Longitudinal Multi-omic Phenotyping Reveals Host-microbe Responses to Bariatric Surgery, Glycaemic Control and Obesity

Nicholas Penney (n.penney@imperial.ac.uk)
Imperial College London https://orcid.org/0000-0003-0896-6590
Derek Yeung
Imperial College London https://orcid.org/0000-0002-5679-9435
Isabel Garcia-Perez
Imperial College London https://orcid.org/0000-0001-8398-0603
Joram POSMA
Imperial College London https://orcid.org/0000-0002-4971-9003
Aleksandra Kopytek
Imperial College London
Bethany Garratt
Imperial College London
Hutan Ashrafian
Imperial College London https://orcid.org/0000-0003-1668-0672
Gary Frost
Imperial University https://orcid.org/0000-0003-0529-6325
Julian Marchesi
Imperial College London
Sanjay Purkayastha
Imperial College Healthcare NHS Trust
Lesley Hoyles
Nottingham Trent University
Ara Darzi
Imperial College London https://orcid.org/0000-0001-7815-7989
Elaine Holmes
Murdoch University

Article

Keywords: Gut Microbiome, Gut Microbiota, Bariatric Surgery, Diabetes, Obesity, Roux-en-Y Gastric Bypass, Sleeve Gastrectomy, metabolomic

Posted Date: December 28th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1146471/v1

Version of Record: A version of this preprint was published at Communications Medicine on October 7th, 2022. See the published version at https://doi.org/10.1038/s43856-022-00185-6.

Longitudinal multi-omic phenotyping reveals host-microbe responses to bariatric surgery, glycaemic control and obesity

Nicholas C Penney[1,2]; Derek KT Yeung[1,2]; Isabel Garcia-Perez[1]; Joram M Posma[1,3]; Aleksandra Kopytek[1]; Bethany Garratt[2]; Hutan Ashrafian[1,2]; Gary Frost[1]; Julian R Marchesi[1]; Sanjay Purkayastha[2]; Lesley Hoyles[1,4]; Ara Darzi[2,5]; Elaine Holmes[1,6]

Author affiliations:

1 Department of Metabolism, Digestion and Reproduction, Faculty of Medicine, Imperial College London, London, SW7 2AZ, UK

2 Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, W2 1NY, UK

3 Health Data Research UK, London, NW1 2BE, UK

4 Department of Biosciences, Nottingham Trent University, Nottingham, NG11 8NS, UK

5 Institute of Global Health Innovation, Imperial College London, London, W2 1NY, UK

6 Centre for Computational & Systems Medicine, Health Futures Institute, Murdoch University, Western Australia, 6150, Australia

Corresponding authors:

Dr Nicholas C. Penney

Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, UK

n.penney@imperial.ac.uk

07545461256

Professor Elaine Holmes (Lead Contact):

Department of Metabolism, Digestion and Reproduction, Faculty of Medicine, Imperial College London, London, UK

elaine.holmes@imperial.ac.uk

Abstract

Resolution of type-2 diabetes (T2D) is common following bariatric surgery, particularly Roux-en-Y gastric bypass (RYGB). However, the underlying mechanisms have not been fully elucidated. To address this we compared the integrated serum, urine and faecal metabolic profiles of obese participants with and without T2D (n=81, T2D=42) with participants who underwent RYGB or sleeve gastrectomy (pre and 3-months post-surgery; n=27), taking diet into account. We co-modelled these data with shotgun metagenomic profiles of the gut microbiota to provide a comprehensive atlas of host-gut microbe responses to bariatric surgery, weight-loss and glycaemic control at the systems level. Bariatric surgery reversed a number of disrupted pathways characteristic of T2D. The differential metabolite set representative of bariatric surgery overlapped with both diabetes (19.3% commonality) and BMI (18.6% commonality). However, the percentage overlap between diabetes and BMI was minimal (4.0% commonality), consistent with weight-independent mechanisms of T2D resolution. The gut microbiota was more strongly correlated to BMI than T2D, although we identified some pathways such as amino acid metabolism that correlated with changes to the gut microbiota and which influence glycaemic control. Improved understanding of GM-host co-metabolism may lead to novel therapies for weight-loss or diabetes.

Keywords

Gut Microbiome, Gut Microbiota, Bariatric Surgery, Diabetes, Obesity, Roux-en-Y Gastric Bypass, Sleeve Gastrectomy, metabolomic

Introduction

The global epidemic in obesity and associated disease states carries a significant health and economic burden. The gut microbiota (GM) has been implicated as a contributing factor in a number of these diseases, including obesity and type-2 diabetes (T2D)[1, 2, 3]. Faecal microbiota transplant experiments in obesity[3] and T2D[3, 4] have shown that this relationship is causal, but these studies have failed to fully unravel the complex mechanisms behind this observation, further complicated by the fact that each individual's GM is unique and subject to redundancy in its metabolic function[5]. Therefore, there is a need to move beyond simply profiling the composition of GM communities in order to understand the true nature of host-microbe relationships.

Surgical procedures such as Roux-en-Y gastric bypass (RYGB) and vertical sleeve gastrectomy (VSG) achieve sustainable weight-loss in obesity[6]. Importantly, they are also highly successful in the resolution of obesity-related co-morbidities including T2D[7]. These metabolic outcomes are achieved through both weight-dependent and, interestingly, weight-independent mechanisms[8]. Weight-independent effects occur because bariatric surgery, particularly RYGB, induces a complex system-wide metabolic effect, including modification of the GM-host metabolic axis[9]. The overwhelming disruption to the GM caused by bariatric surgery is only just being defined, as is its functional importance. To date, few studies have explored longitudinal host-microbe interactions in human cohorts following bariatric surgery, with most studies focussing on either the microbiota or the metabolome[10, 11]. Multiple mechanisms for the GMs contribution to achieving weight-loss and metabolic improvement post-surgery have been hypothesised, including: reduced energy harvest of non-digestible food types such as complex carbohydrates; reduced gut permeability leading to decreased systemic inflammation; and alterations in microbe-host co-metabolites such as bile acids (BAs), amino acids (AAs) and short-chain fatty acids (SCFAs)[10, 12].

Bariatric surgery provides a unique opportunity to unravel these complex host-microbe interactions through longitudinal phenotyping before and after intervention to reduce the impact of inter-individual variability. Here we have performed multi-platform profiling, to establish changes in the host-microbe interactions in volunteers with obesity +/- T2D and in individuals undergoing bariatric surgery with and without T2D to identify dysregulated pathways in T2D that are functionally restored after bariatric surgery. First, we have compared differences in GM-host co-metabolism in participants with T2D compared to non-diabetic individuals at baseline to ascertain which metabolites were

associated with glycaemic control. Next, we profiled subgroups of patients undergoing RYGB or VSG to evaluate changes in GM-host co-metabolism following these contrasting interventions and assessed their impact on glycaemic control, taking into account intervention-dependent changes in eating behaviour.

Results

Serum samples were collected from 158 participants with obesity. Sixty-six participants had T2D and 26 had Impaired Glucose Tolerance (IGT). Complete sample sets of serum, 24-hour urine and stool were collected from 81 of these individuals (42 T2D, 11 IGT).

Forty-nine patients underwent bariatric surgery (VSG=26, RYGB=23; Figure 2A) and gave serum samples pre and 3-months post-surgery (19 T2D, 6 IGT). Twenty-seven of these participants (VSG=14, RYGB=13) gave complete sample sets of serum, 24-hour urine and stool pre and 3-months post-surgery. More diabetic patients underwent RYGB than VSG (11/23 vs 8/26). Otherwise, baseline demographics were not significantly different between procedures. Full demographics are detailed in the Supplemental Information R1.

Microbial and metabolic profiling indicated systematic differences relating to obesity, T2D and bariatric surgery (both RYGB and VSG), with metabolic signatures identified across the three biofluids (urine, serum, faecal water). Each condition had a specific set of metabolic correlates, with some overlap between groups. We identified 207 metabolites associated with bariatric surgery, 54 (26%) of these metabolites were characteristic of improved glycaemic control, 41 (20%) associated with BMI reduction and 28 (14%) were associated with dietary changes (Figure 1). Consistent with the observation that the mechanism for T2D resolution following bariatric surgery is partially independent of weight-loss, of the 175 metabolites associated with either T2D or BMI only 7 overlapped (4% commonality).

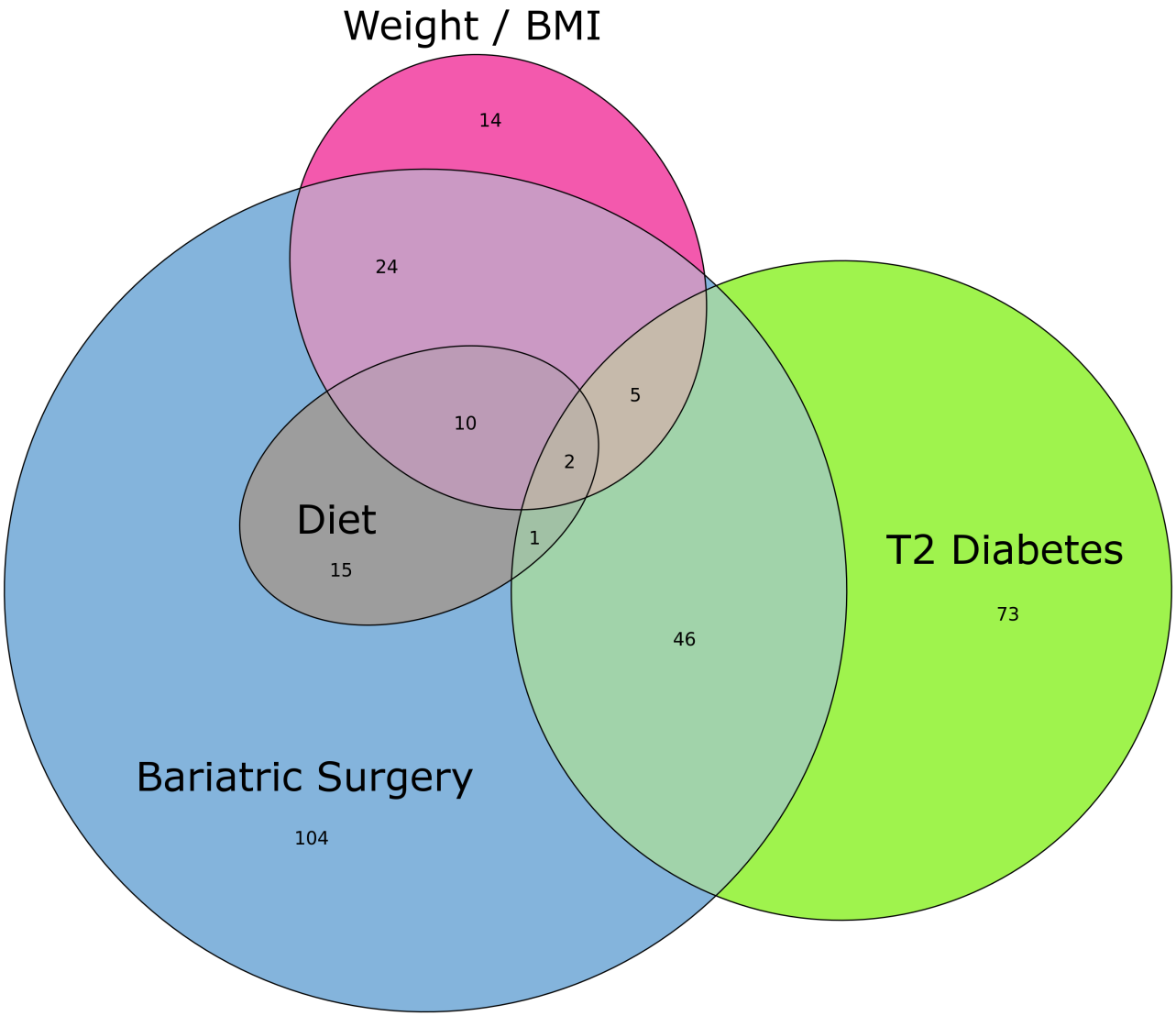

Figure 1. Euler diagram of identified metabolites associated with Bariatric Surgery, Weight / BMI, T2D and Diet.

Euler diagram of identified metabolites from serum, urine and faecal biofluids with associations (pFDR<0.05) to 1) post- versus pre-bariatric surgery, 2) lower weight / BMI, 3) lower HbA1c / non-T2D Vs T2D, 4) lower dietary substrate / higher Alternative Healthy Eating Index (AHEI-2010) score. Metabolites with concordant changes are grouped together. Metabolites from each grouping and their associations are detailed in Supplemental Information R2.

Gut microbial differences between T2D and non-diabetic individuals

Shotgun metagenomic profiling did not identify a difference in microbial gene richness or β-diversity of the GM derived from the distal colon between T2D and non-diabetic participants at baseline. However, compositional analysis of the GM demonstrated lower relative abundance of the genera *Escherichia* (*Proteobacteria*), *Peptostreptococcaceae* (*Firmicutes*) and *Barnesiella* (*Bacteroidetes*) in T2D relative to non-diabetic controls. Individual species from the genera *Ruminococcus*, *Parabacteroides* and *Bacteroides* had higher relative abundance (Figure 3D). Functional analysis of the GM found lower levels of KEGG pathways relating to cofactor and vitamin metabolism, including nicotinate and nicotinamide metabolism and one-carbon metabolism by folate in T2D relative to non-diabetic controls. Methane metabolism, streptomycin and neomycin biosynthesis, polycyclic aromatic hydrocarbon degradation and *D*-alanine metabolism pathways were also less prevalent in T2D, while bisphenol degradation pathways were higher.

Metabolic differences between T2D and non-diabetic individuals

Individuals with T2D were metabolically distinct from non-diabetic participants as determined from both targeted MS assays and global ^{1}H-NMR profiles. In participants with T2D, the serum secondary to primary (2:1) BA ratio was higher compared to non-diabetics at baseline. Conjugation of the primary BA cholic acid (CA), the overall glycine:taurine conjugation ratio of primary BAs and lithocholic acid were also higher in T2D. Conversely, CA and the CA:CDCA (chenodeoxycholic acid) ratio were lower. In addition, conjugation of secondary BAs including tauro-ursodeoxycholic acid (TUDCA), taurohyocholic acid (THCA), glycohyocholic acid (GHCA) and the conjugated UDCA:UDCA ratio was lower. In faeces, 5α-cholanic acid-3α-ol-6-one was lower in individuals with T2D relative to non-diabetic controls. Targeted GC-MS analysis found higher 2-hydroxybutyrate and lactate in T2D participants' urine and serum relative to controls. Serum 2-methylbutyrate and isovalerate were also higher in T2D participants, while urinary butyrate was lower. Further quantitative analyses of serum metabolites found higher branched-chain AAs (BCAAs) leucine, isoleucine and valine, aromatic AAs (AAAs) phenylalanine and tyrosine, as well as alanine, methionine, glutamate, lysine and proline in volunteers with T2D relative to non-diabetic controls. The biogenic amines 2-aminoadipate, methionine sulfoxide and sarcosine and short-chain acylcarnitines (C2-5, 9) were also present in higher concentrations. Conversely, longer-chain acylcarnitines (C14, 16, 18) and a number of lyso-

phosphatidylcholines, acyl-alkyl-phosphatidylcholines, longer-chain diacyl-phosphatidylcholines and sphingomyelins were present in lower concentrations in T2D.

Orthogonal Partial Least Squares – Discriminant Analysis (OPLS-DA) models of global ^{1}H-NMR spectra comparing those with and without T2D were generated. Serum and faeces produced robust models (R^2Y=0·61, Q^2Y=0·44 and R^2Y=0·59, Q^2Y=0·35 respectively) while urine produced the least robust model (R^2Y=0·80, Q^2Y=0·12). In serum, participants with T2D were characterised by higher concentrations of VLDL/LDL lipoproteins, BCAAs, lactate, alanine, proline, pyruvate, tyrosine and α-glucose relative to non-diabetic controls, whereas HDL, glutamine, glycerophosphocholine, phosphocholine/choline and histidine were lower. In faeces, individuals with T2D had higher levels of glycine and the anti-hyperglycaemia drug dimethylbiguanide (Metformin). Lactate, uracil, BCAAs, and tyrosine were lower relative to non-diabetic controls. As expected, in urine, α- and β-glucose were higher in T2D participants, while isobutyrate, glycine, creatine, creatinine, *O*-acetylcarnitine, *N*-methyl-2-pyridone-5-carboxamide, methylnicotinamide and formate were lower relative to non-diabetic controls.

Integrative analysis of metabolic and gut microbiota profiles in T2D and non-diabetic individuals

Multi-omic signatures of participants with T2D versus non-diabetic individuals were modelled using Data Integration Analysis for Biomarker discovery using Latent cOmponents (DIABLO), with a cross-validated balanced error rate (BER) of 0.18, indicating good class separation at the systems level (see Supplemental Information R4). Volunteers with T2D were characterised by higher levels of lactate, glucose and alanine in serum; lactate, glucose and 2-hydroxybutyrate in urine and dimethylbiguanide (Metformin) in faeces compared to non-diabetic controls. Whereas lower serum phosphatidylcholines and HDL lipoprotein; lower urinary glycine, trimethylamine and isobutyrate and faecal valine and uracil levels were characteristic of T2D participants relative to non-diabetics. Lower levels of GM from the *Peptostreptococcaceae* (*Firmicutes*) and *Barnesiella* (*Bacteroidetes*) genera were also seen in the T2D participant signature, as were lower levels of GM KEGG pathways including lipid and *N*-glycan biosynthesis; nicotinate, nicotinamide, methane, alanine and one-carbon metabolism; and biosynthesis of secondary metabolites such as streptomycin and neomycin.

Clinical findings after bariatric surgery

Significant weight-loss was achieved after both procedures (Figure 2B) although the percent weight-loss was greater after RYGB (p=0·023) (Figure 2B). At 3-months post-surgery T2D patients who underwent VSG and RYGB had mean (± standard deviation) glycated haemoglobin (HbA1c) reductions of 17·8mmol/mol (±11·2) and 19·4mmol/mol (±12·5) respectively. Relative to baseline, VSG and RYGB patients had reductions in HbA1c of 27·6% (±12·1) and 28·6% (±14·2) respectively. Three months after VSG, 4/8 T2D participants had complete diabetes resolution (HbA1c <42mmol/mol), 2/8 had partial resolution (HbA1c <48mmol/mol) and 2/8 had ongoing T2D (HbA1c >48mmol/mol). Of 4 participants with IGT that underwent VSG, two had resolution. Three months after RYGB, 8/11 T2D participants had complete diabetes resolution and 3/11 had ongoing T2D. Two participants with IGT underwent RYGB, both had complete resolution.

Three months following surgery there were large changes in dietary intake compared to baseline. As measured by three self-reported 24-hour dietary recall questionnaires collected at each time point, there were significant reported reductions in calorie, carbohydrate, fat, fibre and sodium intake after both operations. In addition, protein intake was significantly reduced after VSG, while saturated fat and sugar intake were significantly reduced after RYGB (Figure 2C). Dietary healthiness, measured using the Alternative Healthy Eating Index (AHEI-2010)[13], was increased 3-months after RYGB and VSG surgery (median 42 vs 54, p=0·002, Figure 2D).

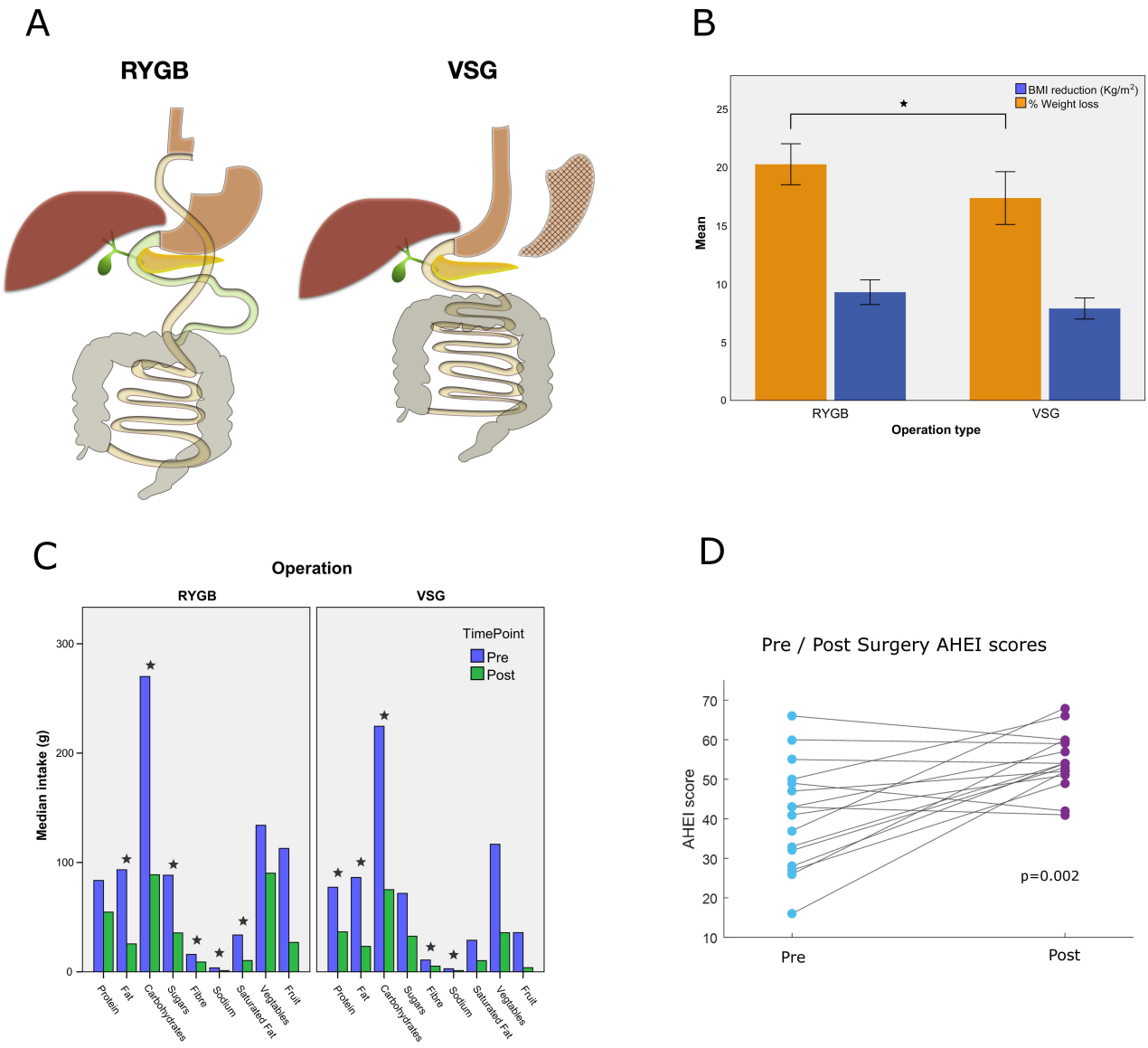

Figure 2. Clinical findings after bariatric surgery.

*(A) Schematic of Roux-en-Y Gastric Bypass (RYGB) and Vertical Sleeve Gastrectomy (VSG) procedures, showing the respective anatomical changes. (B) Weight-loss 3-months post RYGB (n=23) and VSG (n=26) procedures. Patients who underwent VSG had a mean (SD) weight-loss, percent weight-loss and BMI reduction of 21·7kg (6·8), 17·3% (5·6) and 7·9kg/m2 (2·2) respectively. Patients who underwent RYGB had a mean (SD) weight-loss, percent weight-loss and BMI reduction of 24·6kg (7·0), 20·2% (4·1) and 9·3kg/m^2 (2·4) respectively. 95% confidence intervals are shown. Percent weight-loss was significantly greater in the RYGB group (p=0·02). (C) Dietary changes 3-months post RYGB and VSG procedures. (D) Dietary healthiness measured by the Alternative Healthy Eating Index (AHEI) pre and 3-months post bariatric procedures. * = p<0·05.*

Gut microbial changes after bariatric surgery

Multivariate analysis demonstrated a change in β-diversity (Bray-Curtis dissimilarity) after RYGB (PERMANOVA p=0·002), but not VSG (Supplemental Information R3). There was no difference in microbial gene richness after either bariatric procedure. Compositional analysis revealed a major disruption to the GM after RYGB, but more subtle changes after VSG. Three months after RYGB, participants had: i) Increased relative abundance of *Veillonella*, *Gemella*, *Granulicatella*, *Enterococcus*, *Streptococcus* and *Clostridium* (Firmicutes), *Fusobacterium* (Fusobacteria), *Klebsiella* and *Escherichia* (Proteobacteria), *Actinomyces* and *Anaerotruncus* (Actinobacteria) and *Prevotella* (Bacteroidetes); ii) decreased relative abundance of *Holdemania*, *Eubacterium*, *Faecalibacterium*, *Subdoligranulum*, *Ruminococcus*, *Dorea* and *Anaerostipes* (Firmicutes), *Burkholderiales* (Proteobacteria), *Bifidobacterium* and *Collinsella* (Actinobacteria). Changes at each taxonomic rank from phylum to species are shown in Figure 3A. Functional analysis found an increase in bacterial KEGG pathways pertaining to: AA metabolism, lipid metabolism including fatty acid degradation, α-linolenic acid metabolism; and xenobiotic biodegradation including benzoate, aminobenzoate and ethylbenzene degradation. Bile salt hydrolase (choloylglycine hydrolase) pathways were reduced (p=0·03) (Figure 3B).

Three months after VSG, participants had increased relative abundance of a select number of species within the genera *Streptococcus*, *Eubacterium* and *Anaerotruncus* (Firmicutes) and *Escherichia* (Proteobacteria) and decreased species within *Faecalibacterium*, *Dorea*, *Anaerostipes, Roseburia* and *Coprococcus* (Firmicutes) (Figure 3C). Limited changes to KEGG pathways were found after VSG.

Figure 3. Gut microbiota changes after bariatric surgery and in T2D vs non-diabetic controls.

(A) Phylogenetic tree of significant differentially abundant taxa from phyla to species 3-months post RYGB. (B) Phylogenetic tree of significant differentially abundant level 2 and level 3 KEGG pathways 3-months post RYGB. (C) Phylogenetic tree of significant differentially abundant taxa from phyla to species 3-months post VSG. (D) Phylogenetic tree of significant differentially abundant taxa from phyla to species between T2D and non-diabetic participants. Taxa / KEGG pathways significantly increased post-surgery (A-C) or lower in T2D vs non-diabetic (D) are shown in green, taxa / KEGG pathways significantly decreased post-surgery (A-C) or higher in T2D vs non-diabetic (D) are shown in red. Changes that remain significant after Benjamini-Hochberg multiple testing corrections are denoted with an asterisk (*). Plots A-D are shown individually in Supplemental Information R3.

Metabolic changes after bariatric surgery

After RYGB and VSG, both groups had increased secondary (UDCA) and conjugated secondary BAs (GUDCA+TUDCA) in serum. Increases in glycine conjugation relative to taurine in both primary and secondary serum BAs occurred after RYGB. Whereas decreased primary and conjugated primary BAs were noted after VSG. Additionally, GHCA was increased after RYGB, while murocholic acid and isolithocholic acid were increased after VSG. Faecal levels of primary, secondary, conjugated primary and secondary and overall faecal BAs were decreased after both procedures. Similar to serum there was an increase in the conversion ratio of primary to secondary BAs in faeces. Absolute 5α-cholanic acid-3α-ol-6-one and allolithocholic acid were increased after RYGB and VSG respectively. 5β-cholanic acid-3α-ol-12-one and 3,6/3,12-diketocholanic acid were reduced after RYGB. After RYGB, but not VSG, excretion of 2-methylbutyrate and isovalerate in urine and faeces increased. In serum, lactate was decreased after RYGB and showed a non-significant trend towards decreasing after VSG, consistent with the lower concentrations of serum lactate in non-diabetics, compared to diabetics. In faeces, acetate was reduced after both procedures, while butyrate and valerate reduced after VSG. Both procedures resulted in a decrease in the majority of serum AAs, including BCAAs and AAAs. However, glycine and serine were increased after both procedures, with glutamine also increased after RYGB. Serum kynurenine and sarcosine were decreased and symmetric dimethylarginine was increased after both procedures. 2-Aminoadipate was lower after VSG. Short-chain acylcarnitines (C0, C3-5) and a large number of phosphatidylcholines and lyso-phosphatidylcholines species were reduced after both RYGB and VSG. Whereas, C2 and longer-chain acylcarnitines (C7, 8, 10, 14, 16, 18), sphingomyelins and predominantly longer-chain acyl-alkyl-phosphatidylcholines were increased, with similar findings after both interventions.

Repeated Measures, Monte-Carlo Cross-Validation, Partial Least Squares Discriminant Analysis (RM-MCCV-PLSDA) models of serum and urine global ^{1}H-NMR spectra found excellent separation between participants pre- and post-surgery, with robust models after both procedures (Figure 4). In serum, a number of significant changes were consistent after both RYGB and VSG, including altered AA metabolism, matching results in the targeted AA analysis. Ketone bodies acetone, acetoacetate and 3-hydroxybutyrate were increased and evidence of changes to the tricarboxylic acid cycle was seen, with citrate increasing and pyruvate decreasing after surgery. VLDL/LDL, lipid glycerol and choline decreased 3-months after RYGB and VSG. In addition, lactate decreased after RYGB, consistent with the quantitative GC-MS analysis. In urine, a number of significant metabolic changes were also

consistent after both RYGB and VSG procedures, including increased bacterially derived metabolites such as phenylacetylglutamine (PAG), *N*-methyl-4-pyridone-3-carboxamide, 4-cresylsulfate, hippurate, trimethylamine-*N*-oxide (TMAO) and 2-aminobutyrate. Decreased urinary excretion of lactate, 3-hydroxyisovalerate and AAs occurred and a reduction in analgaesic use was observed. RM-MCCV-PLSDA models of faecal samples analysed by ^{1}H-NMR pre- and post-surgery produced a robust model for VSG but not RYGB (Supplemental Information R2). Tyramine and β-alanine were increased after VSG. In keeping with the quantitative SCFA analysis, acetate, butyrate, valerate were decreased, as were isovalerate, lactate, methanol, formate, trimethylamine and phenylacetate.

Full details of metabolic changes can be found in Supplemental Information R2.

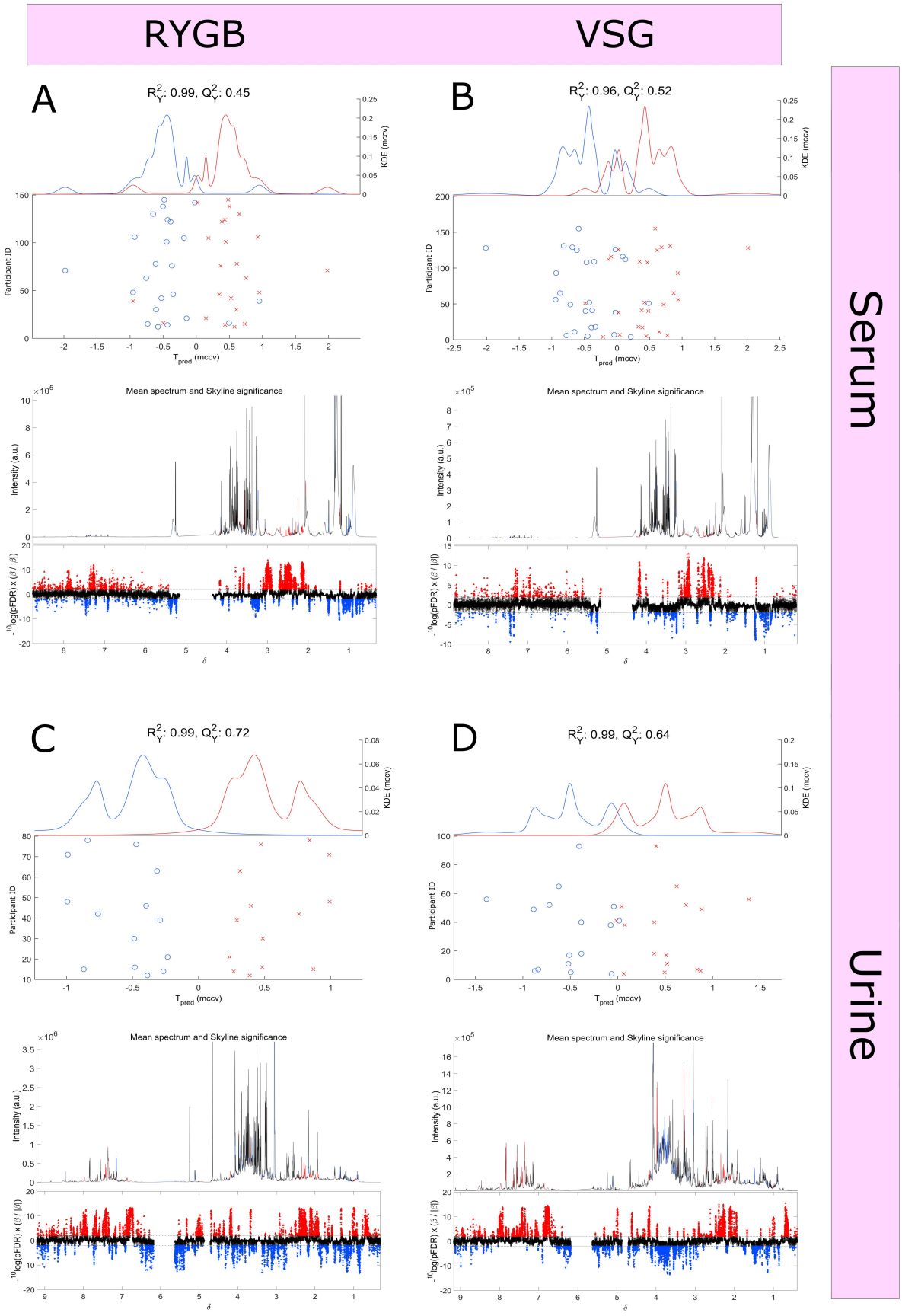
RYGB
VSG
Serum
Urine
A
B
C
D
R^2_Y: 0.99, Q^2_Y: 0.45
R^2_Y: 0.96, Q^2_Y: 0.52
R^2_Y: 0.99, Q^2_Y: 0.72
R^2_Y: 0.99, Q^2_Y: 0.64
KDE (mccv)
Participant ID
T_{pred} (mccv)
Mean spectrum and Skyline significance
Intensity (a.u.)
$-^{10}\log(pFDR) \times (\beta / |\beta|)$
δ
Pre-Surgery
Post-Surgery

Figure 4. Metabolite changes after RYGB and VSG procedures, assessed by ^{1}H-NMR.

RM-MCCV-PLSDA models were generated using ^{1}H-NMR spectra derived from serum and urine biofluids pre and 3-months post RYGB and VSG procedures. Model scores: (A) Serum RYGB (n=23, R^2Y 0·99, Q^2Y 0·45), (B) Serum VSG (n=26, R^2Y 0·96, Q^2Y 0·52), (C) Urine RYGB (n=14, R^2Y 0·99, Q^2Y 0·72), (D) Urine VSG (n=16, R^2Y 0·99, Q^2Y 0·64), (Supplemental Information R2) Faeces RYGB (n=10, R^2Y 0·94, Q^2Y 0·06), (Supplemental Information R2) Faeces VSG (n=14, R^2Y 0·98, Q^2Y 0·66). Upper panels: RM-MCCV-PLSDA scores plots comparing participant samples pre and 3-months post bariatric surgery. Models are comprised of 1 predictive and 1 orthogonal component. Lower panels: Mean ^{1}H-NMR spectrum and Manhattan plot. Manhattan plot showing -log_{10}(pFDR) x sign of the variable regression coefficient for each variable within the RM-MCCV-PLSDA model. Dotted lines illustrate the pFDR significance cut off level (0·01) on the log_{10} scale. Spectra considered significant are highlighted in the Manhattan plot and mean spectrum. Red metabolites (^{1}H-NMR signals) are significantly increased post-surgery, blue metabolites (^{1}H-NMR signals) are significantly decreased post-surgery.

Integrative analysis of metabolic and gut microbiota profiles after bariatric surgery

Multi-omic signatures of the response to RYGB and VSG were modelled using multilevel DIABLO, with cross-validated balanced error rates (BER) of 0·021 and 0·085 respectively, indicating significant differences at the systems level in response to RYGB and VSG and excellent classification between pre- and post-surgery states (Figure 5A-F). Each procedure had a distinct signature. Notably, RYGB was characterised by an increase in tyrosine and phenylalanine metabolism and benzoate and fatty acid degradation pathways in gut bacteria. This difference corresponded with increases in urinary metabolites of bacterial origin including PAG, hippurate and TMAO. In addition, there was increased serum glycine and glycine conjugation of CA and decreased tryptophan and valerate. A number of BA changes were seen in faeces including increased glyco-ursocholanic acid and 5β-cholanic acid 3α-ol-6-one. VSG was characterised by a decrease in a number of GM KEGG pathways. These correlated with decreased faecal SCFA levels, specifically acetate, butyrate and valerate. However, similar to RYGB a number of bacterially derived compounds were increased in the urine including indoxylsulfate, 4-cresylsulfate, hippurate and PAG. Serum changes consisted predominantly of decreased phosphatidycholines.

A multi-omic signature differentiating between the two bariatric procedures was characterised by a greater increase in urinary compounds of microbial origin including PAG, indoxylsulfate, TMAO and 4-hydroxybutyrate after RYGB relative to VSG. The reduction in faecal acetate, butyrate, valerate was greater after VSG, while an increase in 2-methylbutyrate and isovalerate was specific to RYGB. Glycine conjugation of serum BAs was greater after RYGB. A number of bacterial species increased

further after RYGB, including *Escherichia coli* and unclassified *Granulicatella* and *Gemella* species, as did GM KEGG pathways including tryptophan metabolism, benzoate and toluene degradation and biosynthesis of unsaturated fatty acids (Supplemental Information R4).

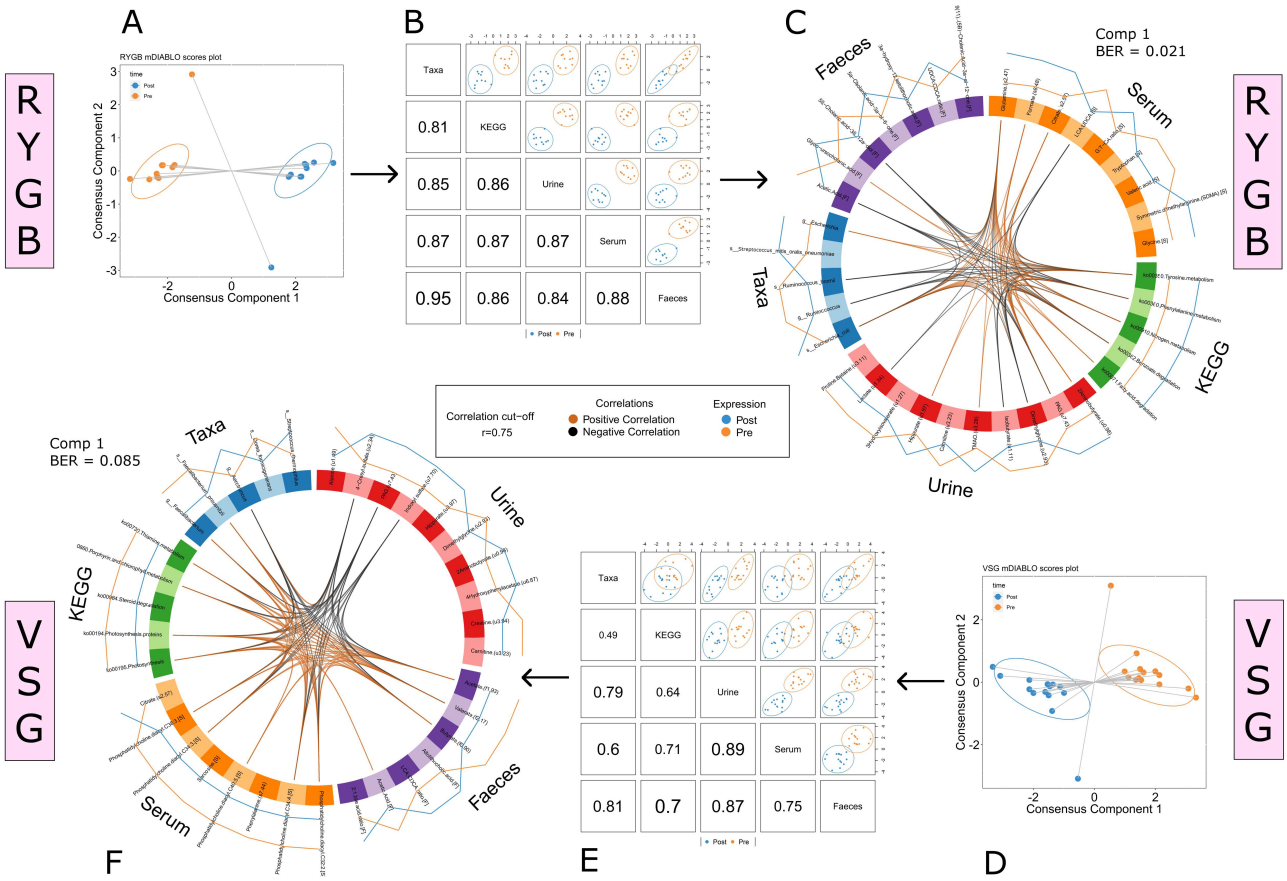

Figure 5. Integration of metabolite and gut microbiota datasets.

Multi-omic datasets pre and 3-months post RYGB (A-C) and VSG (D-F) surgery were integrated using multilevel Data Integration Analysis for Biomarker discovery using Latent cOmponents (mDIABLO). (A, D) Scores plots with samples projected in latent space for RYGB (n=10) and VSG (n=14) models respectively. Classes (pre-/post-surgery) are discriminated along component 1. (B, E) Scores plots derived from component 1 of individual datasets, showing correlations between variables from each dataset. (C, F) Variables from component 1 (coloured by dataset) discriminating between pre/post timepoints are displayed. Differences in expression of variables pre-/post-surgery are shown by the colour-coded outer lines. Correlations (r>0·75) between variables are shown by the colour-coded inner lines. Cross-validated balanced error rates (BER) for each model are shown. Ellipses correspond to 95% confidence intervals.

Integrative correlation analyses of BMI and HbA1c

Body Mass Index (BMI) is closely associated with glycaemic control (HbA1c) and both decreased following bariatric surgery. However, interestingly, BMI and HbA1c correlation networks had contrasting compositions (Figure 6). BMI was strongly correlated to a number of bacterially derived factors. Species including *Escherichia coli*, *Streptococcus anginosus*, *Streptococcus parasanguinis*, *Clostridium hathewayi* and multiple *Veillonella* species were correlated with a lean phenotype, while *Eubacterium rectale* was correlated with increased adiposity. Subsequently a number of GM KEGG pathways such as tryptophan metabolism, linoleic acid metabolism, cytochrome P450, steroid biosynthesis and xenobiotic degradation were also correlated with lower BMI. Bacterially derived urinary metabolites PAG, 4-cresylsulfate, indoxylsulfate and 4-hydroxyphenylacetate were negatively correlated with BMI. While SCFAs acetate and valerate in faeces and isobutyrate in urine correlated positively, as did a number of BA species.

Conversely, HbA1c correlated positively with a range of serum AAs including BCAAs, acylcarnitines (C3, C5, C5-OH), lactate, kynurenine and 2-aminoadipate. In validation of the model, HbA1c increased with higher serum and urinary glucose, faecal metformin levels and age, as expected. Whereas serum HDL, sphingomyelins, glycerophophocholine and GHCA, urinary 2-aminobutyrate and glycine and faecal 5α-cholanic acid 3α-ol-6-one levels were negatively associated with HbA1c.

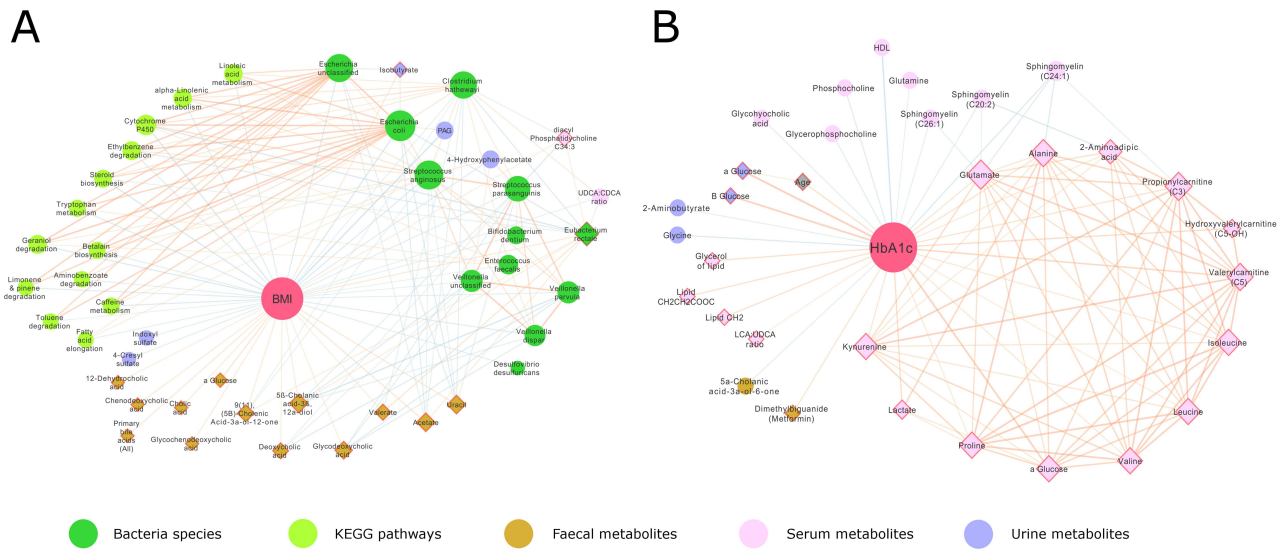

Figure 6. Metabolite and gut microbiota correlations with BMI and HbA1c.

First-order Spearman's correlations to BMI (A) and HbA1c (B) derived from metabolite and microbiota datasets. Correlations with a corrected pFDR <0·01 are shown. Variables with a negative correlation to BMI (A) and HbA1c (B) are displayed within a circle, variables with a positive correlation have a diamond shape and red outline. Node size is proportional to the number of significant correlations to that variable. Positive correlations between variables are shown with orange lines, negative correlations have blue lines. Line thickness is proportional to the correlation strength (r).

Influence of diet on gut microbial and metabolic profiles

Metabolic changes associated with diet overlapped principally with the effects of bariatric surgery and BMI, but not T2D (Figure 1). Carbohydrate and calorie intake correlated positively with a number of serum glycerophospholipids and sphingomyelins as well as phenylalanine, hydroxypropionylcarnitine (C3-OH), propenoylcarnitine (C3:1) and faecal acetate levels (see Supplemental Information R4). While urinary PAG, 4-cresylsulfate and serum acetate were negatively correlated with carbohydrates and calorie intake. PAG was also negatively correlated with fat intake, including trans and polyunsaturated fats. Serum glutamate was higher with higher levels of trans fat and tauro-ursodeoxycholic acid (TUDCA) in serum was negatively correlated with salt (sodium) intake. Fibre intake was positively correlated with serum proline and C24:1-OH sphingomyelin levels. While serum leucine levels were lower as diet healthiness (AHEI-2010 score) increased. Limited changes to the GM relating to diet were identified, but percentage carbohydrate intake correlated negatively with the *Proteobacteria* phylum (Supplemental Information R4).

Discussion

Bariatric surgery has been shown to reverse some of the deleterious effects to the multiple organ systems and metabolic pathways that are disrupted in T2D. However, the mechanisms behind the resolution of T2D are not well understood. It is thought that a complex interplay of weight dependent and independent factors[14] are involved including: reduced adipose tissue leading to reduced inflammation and improved insulin sensitivity[15]; changes in the architecture of pancreatic islets such as increased beta cell mass[16, 17]; changes in energy homeostasis and mitochondrial function[18]; restoration of bile acid levels and their impact on the farnesoid-X receptor (FXR) and the transmembrane G protein-coupled receptor 5 (TGR5)[19, 20]; enhanced release of gut hormones (e.g. GLP1[21], ghrelin[22]); reversal of reduced enteroplasticity[23]; and via modulation of the gut microbiome, which is known to contribute to BMI and other risk factors for T2D[24]. In addition, bariatric surgery

has been shown to normalise gene expression that is dysregulated in T2D, particularly mitochondrial genes and those involved with aerobic metabolism[25]. Here we show the contrasting impact of T2D and bariatric surgery on multiple pathways such as aromatic amino acid metabolism (tyrosine), branched chain amino acid metabolism (valine, leucine, isoleucine), one-carbon metabolism (methionine, sarcosine), short chain acyl carnitines (C3, C4, C5:1) and anaerobic glycolysis (lactate, pyruvate), which were higher in T2D and reduced following bariatric surgery. Similarly, other amino acids (glutamine, histidine), bile acids (GHCA, THCA, GUDCA) and lipids (HDL, acyl-alkyl-phosphatidylcholines C32:1, C38:5, C40:5, C42:5, C44:5, C44:6, Sphingomyelins C16:0, C18:0, C18:1, C20:2, C24:1, C26:1) were lower in T2D and increased following bariatric surgery.

The associations of these metabolites with BMI and glycaemic control are described below. However, although the differential metabolite set representative of bariatric surgery overlapped with both diabetes (19.3% commonality) and BMI (18.6% commonality), the percentage overlap between diabetes and BMI was minimal (4.0% commonality). Thus, it appears that the change in metabolism with respect to resolution of T2D is at least in part independent of BMI reduction, consistent with the observation that BMI and HbA1c were differentially associated with metabolic profiles (Figures 1 & 6).

As previously described[26] the clinical and metabolic impact of RYGB was greater than VSG. The RYGB procedure was strongly associated with altered functionality of the gut microbiome as reflected in the urine, serum and faecal metabolomes and in the KEGG pathways associated with the altered microbiome. The microbiome associated with T2D was distinct from the characteristic non-diabetic microbiome but in general the perturbation caused by T2D was of a lower magnitude than the changes in microbial structure and function observed post bariatric surgery. The lesser extent of GM differences in T2D/non-T2D participants relative to the effects of bariatric surgery, coupled with the strong correlation of the GM to BMI but not HbA1c (Figure 6), suggests that the GM predominantly effects weight-dependent mechanisms. However, although the impact of T2D on the GM was more subtle, nevertheless we identified changes in several metabolite groups that correlated with GM changes after bariatric surgery. As discussed below, some of these metabolites, such as BCAAs, influence glycaemic control and may be important in weight-independent mechanisms of T2D resolution following bariatric surgery.

Multi-omic comparison of participants with T2D to non-diabetic BMI matched controls

Although we found differences in various individual taxa within the GM of T2D participants compared to non-diabetic BMI matched controls, there was no overall difference detected in β-diversity. Similarly, others have found only moderate differences in individuals with T2D and large inter-individual variability[1, 27]. Some of the differences in this cohort, such as lower *Clostridium bartlettii* (aka *Intestinibacter bartlettii*) levels relative to non-diabetic controls, are in keeping with changes due to metformin use (86% metformin use in T2D group, see Supplemental Information R1)[28, 29]. However, other differences in T2D participants such as lower *Escherichia coli* were apparent despite metformin's well characterised action in increasing its relative abundance[28, 29]. *Escherichia* subsequently increased following both RYGB and VSG surgery.

In contrast, large metabolic dissimilarities occurred in T2D participants compared to non-diabetic controls and we have defined a T2D metabolic signature, characterised by altered branched chain and aromatic AA, one-carbon, acylcarnitine, lipid, BA and SCFA metabolism. The association of lysine and 2-aminoadipate with HbA1c is consistent with the fact that both metabolites have been associated with increased risk of T2D in the PREDIMED and Framingham Offspring Studies[30, 31]. However, other studies seem to indicate that 2-aminoadipate can modulate insulin secretion and reduce the impact of diabetes[32]. These contrasting results may be due to Maillard reactions which are increased in diabetes and other age related diseases[33]. Similarly, inflammation has been shown to promote the conversion of tryptophan to kynurenine and increased circulating kynurenine has been previously associated with HbA1c and diabetes[34, 35]. Likewise, methionine is susceptible to oxidation to methionine sulfoxide, which was increased in serum in the T2D participants. This has been associated with ageing and is consistent with higher oxidative stress in the T2D participants[36].

Multi-omic assessment of changes post RYGB and VSG surgery

Reductions in the intake of all food groups, in particular carbohydrate, fat, sugar and fibre are commonly noted after RYGB and VSG. In addition to the restrictive elements of both procedures through a reduction in stomach size, bariatric procedures are known to reduce hunger and increase satiety through modulation of anorexigenic hormones[37]. Patients also have a lower brain-hedonic response to food[38] and altered food preferences[39]. This change appears to be reflected here with an increase in dietary healthiness, measured using the AHEI-2010 score, after surgery. However, it

should be noted that all self-reported dietary data are subject to misreporting[40]. Although both surgical procedures increased diet healthiness, several microbial and metabolic changes post-surgery were distinct between interventions thereby implying that dietary change was not the main driver in these effects. Nevertheless, a number of metabolic effects due to dietary changes were observed and are discussed in the relevant sections below.

The magnitude of alteration in faecal bacterial composition following bariatric surgery was far greater than the perturbation in the GM that was associated with T2D. This was also consistent with the alteration in microbial metabolites identified in the serum, faeces, and particularly the urine following bariatric surgery (hippurate, PAG, 4-cresyl sulfate, indoxyl sulfate and TMAO) whereas these metabolites did not strongly differentiate non-diabetics from individuals with T2D. GM composition changes along the intestine's length due to a gradient in a number of factors such as nutrient availability, oxygen levels, pH and antimicrobial activity including BA levels[41]. The changes to the GM after RYGB in this cohort represent a shift in the colonic bacteria towards those usually found in higher concentrations in the small bowel[42, 43]. Additionally, representation of a number of obligate anaerobes increased. Whereas fermentative bacteria, usually found in high concentrations in the colon, decreased after RYGB[43, 44]. As expected, due to the fact that VSG maintains continuity of the gastrointestinal tract, the bacterial changes following VSG were more subtle. Nevertheless, several species increased after VSG from genera that increased after RYGB. Initial studies investigating obesity and the gut microbiota found a higher *Firmicutes*/*Bacteroidetes* ratio in obese individuals[2, 45], while other studies have reported contrasting results, finding the opposite changes or no difference[46, 47]. We found no significant change after either procedure, supporting the notion that the picture is more complex than changes at a phylum level and that changes at lower taxonomic levels and at a functional level may be more relevant[48]. Functional analysis of our cohort revealed a significant shift towards increased bacterial proteolytic fermentation (putrefaction) pathways after RYGB. This increase occurred despite a decrease in dietary protein consumption, although it is possible that malabsorption resulting from altered small-intestine anatomy leads to higher AA concentrations reaching the colon. The change was corroborated by an increase in a number of bacterially derived metabolites, generated through the fermentation of AAs. These included urinary PAG, indoxylsulfate and 4-cresylsulfate derived from the bacterial metabolism of phenylalanine, tryptophan and tyrosine respectively. Each of these amino acid – microbial metabolite pairs had a significant inverse correlation, consistent with upregulated fermentation of proteins and AAs.

Isovalerate and 2-methylbutyrate, derived from bacterial-degradation of BCAAs, were also increased in faeces and urine after RYGB. Urinary PAG has previously been associated with a lean phenotype[49] and was found to be higher after RYGB in a rat model[50]. In this study, PAG was also correlated with a healthier diet (lower carbohydrate and fat intake, including trans and polyunsaturated fats). Similarly, the increase in hippurate, the glycine conjugate of benzoate, post-surgery is consistent with a lean phenotype and increased fruit intake, as well as reduced risk of metabolic syndrome independent of diet[51]. Although bacterial pathways relating to proteolytic fermentation were not significantly increased after VSG, a similar but less pronounced increase in related metabolites also occurred, suggesting that these pathways were also functionally increased after VSG.

Perturbations in AA homeostasis, particularly BCAA and AAAs are associated with insulin resistance[52] and future risk of diabetes[53]. In this study we identified a higher Fischer ratio (BCAA/AAA), previously associated with worsening liver function, in participants with T2D relative to non-diabetics[54, 55]. BCAA and AAAs both reduced following surgery, but the Fischer ratio only reduced following surgery in participants with T2D and impaired glucose tolerance and was not significantly altered following surgery in non-diabetics. The association of BCAA with diabetes is thought to be causative[56]. BCAA infusions in rats and humans leads to the development of insulin resistance[57]. This led us to conclude that the observed changes in AA profiles after surgery are likely to be an important factor driving improved insulin resistance. Our findings suggest that the changes in AA profiles resulted in part from increased AA putrefaction by the GM after RYGB. Consequently, RYGB may confer a metabolic advantage for patients with T2D compared to VSG, through decreased availability and absorption of BCAA from the gut. In support of this, germ-free mice have significantly altered profiles of AAs absorbed from the gut via the portal vein relative to those with a normal GM[58]. This difference is due to the large number of bacteria involved in AA biosynthesis and fermentation[59, 60]. In a prior study, oral gavage of mice with *Bacteroides thetaiotaomicron* led to a reduction in glutamate, phenylalanine, leucine and valine[10]. Here two species of *Paraprevotella,* highly capable of producing AAs, correlated with serum alanine levels. Perhaps more importantly we also identified a number of negative correlations between bacterial species and serum AA levels, indicating that fermentation of AAs by gut bacteria is an important pathway influencing human AA profiles. In particular, species from the genera *Streptococcus* and *Clostridium*, important AA fermenters[60], correlated with a reduction in serum BCAAs (Supplemental Information R4). However, further studies are needed that investigate the combined microbial, dietary and human contribution towards circulating levels of

BCAAs. For example, we also note that the BCAA leucine was negatively correlated with dietary healthiness (AHEI-2010 score) in this cohort.

In keeping with greater changes to the GM following RYGB, urinary TMAO was significantly increased after RYGB but not VSG. Reduction of TMAO and other dietary nutrients such as choline and *L*-carnitine to trimethylamine (TMA) is performed by the GM (predominantly *Enterobacteriaceae*). Absorbed TMA is then converted back to TMAO by host hepatic enzymes[61]. TMAO has been described as pro-atherogenic and high serum levels of TMAO have been proposed to be a predictor of cardiovascular disease[53, 62]. However, the causative effect of cardiovascular disease from high TMAO is disputed and these results may have been due to confounders such as reduced kidney function and poor metabolic control[63]. Other evidence such as the presence of high concentrations of TMAO in elite athletes, as well as in the urine of Japanese populations whose diet contains a high portion of fish and who have a low risk of cardiovascular disease, suggests that the causal role of TMAO in cardiovascular disease is complex and conditional on a wealth of host-microbiota factors[64, 65]. Moreover, TMAO was recently found to protect against impaired glucose tolerance and reduce endoplasmic reticulum stress[66].

The initial gateway step in biotransformation of primary conjugated BAs to secondary BAs is performed by gut bacteria containing the bile salt hydrolase (BSH) enzyme. Interestingly, although we observed increased secondary BAs, we found that the BSH gene load was decreased after RYGB with no overall change after VSG. Furthermore, there were no significant changes in other BA enzymes such as hydroxysteroid dehydrogenases. We concluded that the profound changes in the BA pool composition observed after both procedures, including increased secondary BAs, result predominantly from changes in host factors such as altered hepatic processing and BA re-absorption. In addition, there was a consistent increase in glycine conjugation of BAs relative to taurine in serum after RYGB. Changes in glycine/taurine conjugation can result from changes in their bioavailability in the liver[67] and certainly there was increased serum glycine in this cohort. In addition, bacterial glycine metabolism pathways were increased after RYGB and this may also influence the bioavailability of glycine for conjugation. Indeed, germ-free animals excrete almost exclusively tauro-conjugated BAs[68]. Interestingly, GHCA had a strong negative correlation with HbA1c in this cohort. Diabetes has been associated with lower serum concentrations of HCA species and they are strong predictors of metabolic disease[69]. Administration of hyocholic acid by others increased serum fasting GLP-1 in healthy and diabetic mouse models by simultaneously activating TGR5 and inhibiting FXR, a unique

mechanism not found in other BA species[70]. However, the overall metabolic effects of differences in the BA pool and conjugation patterns are difficult to predict[20].

SCFAs are produced through the bacterial fermentation of dietary fibre and complex carbohydrates. Faecal acetate decreased after both procedures, likely due to a significant reduction in dietary substrate. Indeed, faecal acetate was negatively correlated with dietary carbohydrate and calorie intake. However, the reduction in butyrate and valerate seen after VSG was not replicated after RYGB, despite a similar reduction in dietary substrate. This discrepancy suggests that the GM following RYGB is able to produce these SCFA more readily than the microbiota after VSG. The SCFA signature is important as each SCFA has a unique impact on the host. Acetate production leads to a positive feedback loop that increases appetite, induces lipid deposition in liver and skeletal muscle, and increases insulin resistance[71]. Whereas butyrate has a beneficial role in host satiety, insulin resistance and colonocyte health[72].

Crucially, although the overall changes after bariatric surgery appear to be towards a healthier phenotype there may be some negative consequences of surgery. Increased protein metabolism within the gut is usually considered to be harmful[73]. Putrefaction results in the production of toxic compounds such as amines, sulphides and ammonia[74], whilst phenols and indoles, increased after bariatric surgery in this cohort, are reported to be pro-inflammatory and cytotoxic[75, 76].

In conclusion, metabolic changes post-surgery were achieved by both weight-dependent and weight-independent processes and we have identified multi-omic signatures specific to obese and T2D states at the systems level, some of which demonstrated contrasting patters when compared with the effects of bariatric surgery. BMI correlated inversely with bacterially derived urinary metabolites such as PAG, 4-cresylsulfate and indoxylsulfate and positively with faecal acetate and valerate, whereas HbA1c showed stronger correlation with serum AAs, acylcarnitines, kynurenine and 2-aminoadipate. Greater functional and taxonomic changes were observed in the GM following RYGB compared to VSG. These microbial changes, particularly after RYGB, appeared to influence the complex relationship between the GM and host metabolism. The abundance of amino acid metabolism pathways within the GM and corresponding metabolites of protein putrefaction increased after RYGB, despite reduced dietary protein intake, and correlated with decreased serum BCAAs. Ultimately, further mechanistic work is needed to better understand these GM–host co-metabolism pathways and to establish their full effects, including those that may be harmful to health.

Methods

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Elaine Holmes (elaine.holmes@imperial.ac.uk).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Metagenomic data has been deposited with GenBank, EMBL and DDBJ databases under the BioProject accession number PRJNA473348.

Further metagenomic and metabolic data have been deposited at Mendeley Data: http://dx.doi.org/10.17632/t76nm3yfzh.1

Experimental model and participant details

Recruitment

Patients referred for consideration of bariatric surgery who were obese (BMI >30kg/m^2), aged ≥18, had failed efforts at lifestyle modification and dieting and were willing to comply with the trial protocol were recruited prospectively. Diabetics (HbA1c >48mmol/mol or treated) and non-diabetics were eligible for recruitment.

Patients who had previously undergone bariatric or major abdominal surgery, were or intended to become pregnant during trial period, or took long-term antibiotics were excluded. Major abdominal surgery included patients who had undergone small or large bowel resection, liver, pancreatic, splenic or stomach surgery, as these could influence the gut microbiota and / or the patient's metabolic state. Patients that had previously had an appendicectomy, cholecystectomy or hernia repair were not excluded.

The study protocol and sample collection instructions were co-developed with patient representatives to help reduce the study burden for patients. To improve patient compliance sample collection occurred at the time of patients' usual NHS appointments preoperatively and at 3-months post-procedure. A small exploratory cohort were also sampled at 1-year post-procedure and are reported in Supplemental Information.

Metabolic surgery

Participants underwent Roux-en-Y Gastric Bypass (RYGB) or Vertical Sleeve Gastrectomy (VSG) surgery at a National Health Service (NHS) University Teaching Hospital in London, UK. A single dose of 1.2g intravenous co-amoxiclav was given during induction of anaesthesia (clindamycin if penicillin allergic).

Regulatory approvals

The study received NHS Research Ethics Committee (15/ES/0026) approval and was registered with ClinicalTrials.gov (NCT02421055).

Clinical, demographic and dietary data collection

Participants were assessed at the above time points for: 1) anthropometric & physiological measurements, 2) demographic details, 3) biochemical parameters including glycated haemoglobin (HbA1c), 4) oral-hypoglycaemic, insulin and other medication use, length of diabetes diagnosis and other co-morbidities.

Diet assessment

An online self-reported 24hr dietary recall questionnaire (www.myfood24.org) was utilised to capture detailed dietary intake information from patients at each of the study time points so that changes in diet following surgery could be accounted for in the analysis. Three 24hr recall questionnaires were completed at each time point. Participants were able to pick from a selection of pictures of corresponding foods to accurately ascertain portion quantities. Collected dietary information was used to calculate Alternative Healthy Eating Index 2010 (AHEI-2010) scores as described previously[13]. In brief, scores of 0-10 were given for 11 components (maximum score 110). High scores were given for a high intake of vegetables, fruit, nuts and legumes, whole grains, long chain omega-3 fats and polyunsaturated fats, moderate intake of alcohol and low intake of sugar, sweetened drinks and fruit juice, red and processed meat, trans-fat and sodium.

Sample collection

Serum (n=158), faecal (n=81) and 24hr urine (n=83) samples were collected preoperatively and at 3-months (serum=49, faecal=27, urine=30) and 1-year postoperatively (exploratory 1-year data reported in Supplemental Information due to low numbers), in a non-fasted state. 24-hour urine samples were collected in sterile containers from 9am until 9am on the day of the study visit. Stool samples were collected using a Faecotainer® collection kit and stored on an ice pack provided, as close as possible to but not more than 6 hours before the study visit. Serum was collected using red top BD Vacutainer® serum tubes (no additive) and processed according to manufacturer guidelines. After collection, samples were aliquoted and stored at -80°C until analysis. Prior to freezing, a separate aliquot of homogenised stool was used to generate faecal water as follows: approximately 10g of stool was added to 4 parts HPLC grade H_2O (g/ml), vortexed at 2850rpm for 15min then centrifuged at 10000g for 15min at 4°C. The resulting supernatant (faecal water) was frozen at -80°C until analysis[77].

Method details

^{1}H-NMR metabolite analysis

Sample preparation

Urine and faecal water samples were prepared for analysis by ^{1}H-NMR spectroscopy as follows: frozen samples (-80°C) were thawed, vortexed and then centrifuged at 1600g for 10min to remove particulates and precipitated proteins. Faecal water supernatant was further filtered through Micro centrifuge filters (0.45µm Nylon, Costar) at 16000g for 15min at 4°C. 540µL of each sample was mixed with 60µL of 1.5M KH_2PO_4 buffer (pH 7.4, 80% D_2O) containing 1 mM of the internal reference standard, 3-(trimethylsilyl)-[2,2,3,3,-$^2H_4$]-propionic acid (TSP) and 2mM sodium azide (NaN_3), as described previously[78].

After thawing, serum samples were centrifuged at 12000g for 5 minutes at 4°C. Subsequently, 300µL of serum was mixed with 300µL of 0.075M NaH_2PO_4 buffer (pH 7.4) containing 0.8mM of the internal reference standard, 3-(trimethylsilyl)-[2,2,3,3,-$^2H_4$]-propionic acid (TSP) and 3.1mM sodium azide (NaN_3), as described previously[78].

^{1}H-NMR spectroscopy

^{1}H-NMR spectroscopy was performed at 300K on Bruker 600MHz (urine and serum) and 800MHz (faecal water) spectrometers (Bruker Biospin) using the following standard one-dimensional pulse sequence: RD – g_{z1} – 90° – t_1 – 90° – t_m – g_{z2} – 90° – ACQ[78]. The relaxation delay (RD) was set at 4s, 90° represents the applied 90° radio frequency pulse, interpulse delay (t_1) was set to an interval of 4µs, mixing time (tm) was 10ms, magnetic field gradients (g_{z1} and g_{z2}) were applied for 1ms and the acquisition period (AQA) was 2.7s. Water suppression was achieved through irradiation of the water signal during RD and t_m. For the urine samples, each spectrum was acquired using 4 dummy scans followed by 32 scans while faecal spectra were acquired using 256 scans and 4 dummy scans and collected into 64K data points. A spectral width of 12,000Hz was used for all the samples. Prior to Fourier transformation, the free induction decays (FIDs) were multiplied by an exponential function corresponding to a line broadening of 0.3Hz. Serum samples were analysed by ^{1}H-NMR using the standard one-dimensional pulse sequence described above and Carr-Purcell-Meiboom-Gill (CPMG) one dimensional pulse sequences. CPMG was used to attenuate broad, interfering peaks from lipids and proteins present in serum. The CPMG pulse sequence had the form RD-90°-(t-180°-t)n-ACQ. The acquisition parameters were set using the same settings as the standard 1D pulse sequence, with the spin-echo delay (t) set at 0.3ms and 128 loops (n) performed. Continuous wave irradiation was applied at the water resonance frequency during the relaxation delay (RD).

Pre-processing

^{1}H-NMR spectra were automatically corrected for phase and baseline distortions and referenced to the TSP singlet at δ 0.0 using TopSpin 3.1 software. Spectra were then digitized into 20K data points at a resolution of 0.0005ppm using an in-house MATLAB R2014a (Mathworks) script. Subsequently, spectral regions corresponding to the internal standard (δ -0.5 to 0.5) and water (δ 4.6 to 5) peaks were removed. In addition, urea (δ 5.4 to 6.3) was removed from the urinary and serum spectra due to its tendency to cross-saturate with the suppressed water resonance. All spectra were normalised using median fold change normalisation using the median spectrum as the reference[79].

Quantitative Bile Acid analysis

Quantitative analysis of 57 bile acids was performed using an established technique[80]. The method was adapted for analysis of bile acids in faecal samples.

Sample preparation

Bile acids were extracted from serum using the following method: 100µL of serum was vortexed with 280µL of MeOH. Samples were centrifuged at 14000g for 15min at 4°C, followed by incubation at -20°C for 20min. Internal standards (16 deuterated bile acids) were added to the supernatant at a final concentration of 50nM.

Bile acids were extracted from faecal samples using the following method. Faecal samples were first freeze-dried. 100mg of freeze-dried material was then placed in microtubes with 1ml of 2:1:1 H_2O: Acetonitrile (ACN): Isopropanol (IPA) and approximately 50mg of 1mm Zirconia beads. This underwent 3x30 seconds bead beating and a Biospec bead beater followed by centrifugation at 16,000g for 20min at 4°C. The supernatant was further filtered through Micro-centrifuge filters (0.45µm Nylon, Costar) at 16000g for 15min at 4°C. To ensure bile acid concentrations were within the dynamic range of the machine extracts were diluted 1:25 and 1:200 prior to analysis using the H_2O:ACN:IPA mix. Internal standard (16 deuterated bile acids) was added to the filtered supernatant at a final concentration of 50nM.

LC-MS machine conditions

BA analysis was performed using an ACQUITY ultra-performance liquid chromatography (UPLC) coupled to a Xevo triple quadrupole (TQ-S) mass spectrometer.

For liquid chromatography, an ACQUITY BEH C8 column (1.7µm, 100 mm × 2.1 mm) was used at an operating temperature of 60°C. The mobile phase solvent A consisted of a 1:10 ACN:H_2O, with 1mM ammonium acetate and pH 4.15 adjusted with acetic acid. Mobile phase solvent B consisted of 1:1 ACN:IPA. The chromatographic gradient was as previously published[80].

Mass spectrometry was performed in negative ionisation mode (ESI-) using the following parameters: capillary voltage 1.5kV, cone voltage 60V, source temperature 150°C, desolvation temperature 600°C, desolvation gas flow 1000 L/hr, and cone gas flow 150L/hr. 57 bile acid species (36 non-conjugated, 12 taurine conjugated, 9 glycine conjugated) were assayed using multiple reaction

monitoring (MRM). The transitions for each bile acid and deuterated internal standard was as previously published[80].

Quantitative analysis of SCFAs and other Carboxylic Acids

A total of five short/medium chain fatty acids, three methyl-branched SCFAs and two hydroxyl carboxylic acids were analysed by GC-MS using a method adapted from Moreau et al[81].

Sample preparation

After defrosting and mixing, 100µL of urine / serum was aliquoted with 500µL of methyl *tert*-butyl ether (MTBE) with 100ppm of internal standard (methyl stearate) and 2µL of HCL. This was vortexed and then shaken for 20min. Following this samples were centrifuged at 10000g for 5min at 4°C. Next, 90µL of the polar phase was placed into a silanised vial and vortexed with 150µL of derivatiser *N*-*tert*-butyldimethylsilyl-*N*-methyltrifluoroacetamide with 1% *tert*-butyldimethylchlorosilane (MTBSTFA + 1% TBDMSCl). This was then incubated for 45min at 60°C before aliquoting into silanised inserts for analysis.

The method was modified to account for higher levels of SCFA in stool. After defrosting, 100mg of stool was aliquoted with 1,000µL of MTBE with 100ppm of internal standard (methyl stearate) and 4µL of HCL. 30µL of the polar phase was mixed with 150µL of derivatiser.

GC-MS machine conditions

Derivatised samples were analysed by GC-MS with a Bruker triple quadrupole (TQ) GC-MS/MS. Helium was used as a carrier gas at a constant flow rate of 1.5 ml/min through the column. The injector temperature was 250°C with a split ratio 1:10. The temperature of the oven was started at 40°C and increased at the rate of 46°C/min to 127°C, 2°C/min to reach 131°C, 30°C/min to reach 160°C, then 50°C/min to reach a final temperature of 300°C. The transfer line to the mass spectrometer was set at 280°C. Targeted analysis of the ten compounds and internal standard was performed in multiple reaction monitoring mode (MRN) using the following settings:

Compound	**Quantifier (m/z)**	**Qualifier (m/z)**	**Collision Energy (eV)**
Acetate	117 → 75	117 → 47	10
Propionate	131 → 75	131 → 47	10

Butyrate / Isobutyrate	145 → 75	145 → 43	10
Valerate / Isovalerate	159 → 75	159 → 57	12
2 Methylbutyrate	159 → 75	159 → 57	12
2 Hydroxybutyrate	147 → 73	147 → 45	20
Caproate	173 → 75	173 → 81	15
Lactate	147 → 73	147 → 45	20
Methyl stearate (IS)	87 → 55	87 → 59	10

Quantitative serum metabolite analysis

Quantitative analysis of other metabolites in serum samples, including amino acids, biogenic amines, acylcarnitines, phosphatidylcholines, lysophosphatidylcholines and sphingolipids was performed using the Biocrates AbsoluteIDQ® p180 kit, according to the manufacturer guidelines[82]. Samples were analysed using flow injection analysis (FIA)-MS/MS and LC-MS/MS for different metabolite groups.

Sample preparation

10µL of serum sample / PBS / calibration / QC and 10µL of the ISTD mix (except in blanks) was added to each well. This was dried for 30min under nitrogen flow. Following this, 50 µL of the derivatization solution was pipetted into each well. The plate was covered and incubated for 20min, then dried for 60min under nitrogen flow. Next, 300µL of extraction solvent was added to each well, shaken for 30min at 450rpm, then centrifuged for 2min at 500g. For the LC-MS/MS 150µL was added to 150µl H_2O. For the FIA, 15µL was added to 750µL of FIA mobile phase. Both plates were shaken for 2min at 600rpm.

Machine conditions

Samples were analysed using a Waters I-Class UHPLC system and Waters Xevo TQ-S tandem mass spectrometer.

For FIA-MS/MS (direct infusion): the FIA mobile phase consisted of Biocrates Solvent I + 290 mL MeOH. A 2min isocratic method was used, starting at 0.15mL/min for 0.1min, gradually decreasing to 0.03mL/min at 1min, increasing to 0.2mL/min at 1.5min, to 0.8mL/min at 1.60 min, and finally decreasing to 0.15mL/min at 1.95 min. MS settings were: capillary voltage 3.2kV, cone voltage 10V,

source offset 50V, source temp 150°C, desolvation temp 620°C, cone gas 150L/H, desolvation gas 1000L/H, collision gas flow 0.15mL/min, probe position 5mm.

For LC-MS/MS: A Waters Acquity UPLC BEHC18 1.7µm 2.1 x 75mm column was used. Mobile phase A: 1000mL H_2O + 2mL formic acid (FA), Mobile phase B: 500mL ACN + 1mL FA. Gradient elution was used; starting at a flow rate of 0.8mL/min with 100% A for 0.45min, then changing in a linear gradient to 85% A at 3.3min, to 30% A at 5.9min, to 100% B at 6.05min, flow then increased in a concave gradient to 0.9 mL/min 100% B by 6.20min, remaining at 0.9mL/min 100% B until 6.42min, before decreasing back in a concave gradient to 0.8mL/min 100% B at 6.52min. The mobile phase was then changed in a concave gradient from 100% B to 100% A between 6.52 and 6.7min and remained at 100% A 0.8 ml/min until 7.3min. MS settings were: capillary voltage 3.9kV, cone voltage 20V, source offset 50V, source temp 150°C, desolvation temp 350°C, cone gas 150L/Hr, desolvation Gas 650L/Hr, collision gas flow 0.15mL/min, probe position 7mm.

Data were processed using targetlynx (Waters) and METIDQ (Biocrates; version Carbon) then exported as a CSV file for statistical analysis.

Faecal metagenomic analysis

DNA extraction

Faecal samples were stored at -80°C prior to analysis. DNA was extracted using the MoBio PowerFaecal® DNA Isolation Kit, according to manufacturer's instructions. In brief, DNA was extracted from two separate 0.25g aliquots of mixed whole stool samples at each analysis point. Samples were homogenised in 2ml bead beating tubes containing garnet beads. Cell lysis of host and microbial cells was facilitated through both mechanical collisions between beads and chemical disruption of cell membranes. The reagent to precipitate non-DNA organic and inorganic material was then applied. Lastly, DNA was captured on a silica spin column, washed and eluted for downstream analysis. Quality control of DNA quality and quantity was assessed using an Agilent 4200 TapeStation.

Shotgun sequencing

Shotgun sequencing was performed using an Illumina HiSeq 4000 with paired-end 150bp reads. Library preparation was undertaken using the NEBNext Ultra II DNA Library Prep Kit. 15 dual index

barcodes (unique at both ends) were custom-designed and ordered from Integrate DNA Technologies (IDT®). Quality control of prepped libraries was performed using the Promega GloMax® and QuantiFluor® dsDNA systems. Each of the "uniquely dual-indexed" libraries were pooled and run on a single lane of the HiSeq4000. A mean of 6.89Gb sequence data was acquired for each of 120 samples (median 6.84, range 3.88 – 14Gb).

Processing of sequence data

There are known lane-swapping issues with the HiSeq 4000, leading to duplication of some sequencing reads. For this reason, fastq files for each sample were subject to de-duplication using FastUniq[83]. Sequencing data were then processed using the Scalable Metagenomics Pipeline (ScaMP)[84], (http://www.imperial.ac.uk/bioinformatics-data-science-group/resources/software/scamp). In brief, raw sequence data were assessed for the presence of adapter sequences and trimmed using Trim Galore! (Babraham Bioinformatics) to remove low-quality bases (Q < 20) from the 3' end of reads and discard trimmed reads shorter than 100 nt. Quality control of trimmed reads was performed using FastQC (Babraham Bioinformatics). Reads that mapped with BWA-MEM[85] to human genome (hg19) were removed from read pairs, as ethical permission is not available for use of human data derived from metagenomes. Remaining reads were assumed to be microbial (bacteria, archaea, virus, fungi, protozoa) and processed further. Trimmed sequence data with human reads removed have been deposited with GenBank, EMBL and DDBJ databases under the BioProject accession number PRJNA473348.

MetaPhlAn 2.6[86, 87] was used to determine the bacterial and archaeal taxonomic composition/abundance for each sample. Metagenome assembly was carried out in two rounds using metaSPAdes 3.11.0[88], with an initial independent assembly carried out for each sample. Unassembled reads were then pooled and subjected to a second round of assembly to improve the representation of low-abundance sequences. Taxa were normalised to relative abundance for downstream analyses.

Ab-initio gene prediction was carried out using MetaGeneMark[89, 90]. The resulting predictions were translated, and the protein sequences clustered using the cluster-fast method of UCLUST[91], with a 95 % identity cut-off. Centroid sequences from each cluster were used to form a non-redundant gene catalogue used for downstream analyses. Gene abundance in each sample was determined by alignment of the reads using BWA-MEM against the gene catalogue, determining the number of

reads mapped to each gene sequence and normalising as described[92]. Functional annotation to KEGG pathways was carried out by mapping centroid sequences to the eggNOG-Mapper database[93] (version 4.5, downloaded on 1 March 2018) using diamond on our in-house server.

Microbial gene richness (MGR) was determined as described previously[92, 94]. Briefly, data were downsized to adjust for sequencing depth and technical variability by randomly selecting 7 million reads mapped to the gene catalogue (of 11,005,136 genes) for each sample and then computing the mean number of genes drawn over 30 random samplings.

Quantification and statistical analysis

Pre-processing

To correct for dilution differences between samples normalisation procedures were applied. Global metabolite (^{1}H-NMR spectra) data sets were corrected for dilution effects using median fold change normalisation[79]. Scaling to unit variance was then applied to serum and urine data sets, while pareto scaling was used for faecal datasets, due to the presence of dominant and variable oligosaccharide resonances. Targeted metabolites measured within urine samples were corrected for dilutional differences using osmolality and creatinine measurements[95]. Metagenome data were expressed as relative abundance. Taxa with low abundance (present in <30% of both subgroups) were excluded from downstream statistical analyses.

Univariate analysis

Due to the non-parametric nature of the results, differences between paired samples pre/post intervention in quantified metabolites and within the gut microbiota were assessed for significance using the Wilcoxon Rank test (two-sided). Differences in non-paired data were assessed using the Mann-Whitney U test (two-sided). *P*-values were adjusted for multiple testing using the Benjamini-Hochberg (BH) False Discovery Rate method (pFDR). Phylogenetic Trees were generated to illustrate significant gut microbiota changes using the GraPhlAn[96] script in Python.

Multivariate analysis

Multivariate statistical analysis of normalised ^{1}H-NMR spectra was performed using SIMCA 15 (Umetrics)[79]. Principal component analysis (PCA) was used to provide an overview of the data.

Orthogonal Partial Least Squares – Discriminant Analysis (OPLS-DA) models were established based on one predictive component and one orthogonal component to discriminate between samples from T2D and non-diabetic participants. Unit variance scaling was applied to ^{1}H-NMR spectral data. The fit and predictability of the models obtained were determined by the R^2Y and Q^2Y values respectively. Significant metabolites differentiating between groups were obtained from ^{1}H-NMR OPLS-DA models after investigating ^{1}H-NMR signals with correlation coefficient values higher than 0.35. Jack-knifed 95% confidence intervals of the coefficients were used to confirm significance of the variables.

Longitudinally paired global metabolic data, pre- and post-intervention, were analysed using Repeated Measures, Monte-Carlo Cross-Validation, PLS-DA (RM-MCCV-PLSDA)[97, 98] using covariate adjusted projection to latent structures in MATLAB. Data were centred and scaled to account for the repeated-measures design. 1000 MCCV models were generated and used to calculate the mean cross-validated predictive component score (T_{pred}) and variance for each sample[98]. The fit and predictability of the models obtained was determined by the R^2X and Q^2Y values respectively. Gaussian kernel density estimates of the T_{pred} in each group were generated for visual interpretation[98]. 25 bootstrap resamplings in each of the 1000 models was used to estimate the variance and mean coefficient for each variable and derive a p value for each variable accordingly[98]. Benjamini-Hochberg false discovery corrections were performed and a variable was considered significant with a false discovery rate value (q) ≤ 0.01. Manhattan plots showing $-\log_{10}(q)$ x sign of the variable regression coefficient for each variable within each RM-MCCV-PLSDA model were generated, with dotted lines added to illustrate the q value significance cut off level on the $\log_{10}$ scale.

Exploration of gut microbiota taxa was performed using Principal coordinates analysis (PCoA) of Bray-Curtis dissimilarity matrices (β-diversity) using the Vegan[99] function in R. Significance of group separation in β-diversity was assessed by permutational multivariate analysis of variance (PERMANOVA). Nested PERMANOVA was used longitudinal analyses to account for the repeated measures design.

Metabolite identification

A combination of data-driven strategies such as such as SubseT Optimization by Reference Matching (STORM)[100] and Statistical TOtal Correlation SpectroscopY (STOCSY)[101] and analytical identification strategies were used to aid structural identification of significant discriminatory metabolites. Specifically, a catalogue of 1D ^{1}H-NMR sequence with water pre-saturation and 2D NMR experiments

such as J-Resolved spectroscopy, ^{1}H-^{1}H TOtal Correlation SpectroscopY (TOCSY), ^{1}H-^{1}H COrrelation SpectroscopY (COSY), ^{1}H-^{13}C Hetero-nuclear Single Quantum Coherence (HSQC) and ^{1}H-^{13}C Hetero-nuclear Multiple-Bond Correlation (HMBC) spectroscopy were performed. Finally, when possible metabolites were confirmed by in situ spiking experiments using authentic chemical standards.

Relative concentrations of identified metabolites from ^{1}H-NMR datasets were calculated from intensity measurements of a representative spectral peak of the metabolite, ensuring no overlap with signals from other metabolites.

Euler diagram of metabolites

A Euler diagram of identified metabolites from Serum, Urine and Faecal biofluids associated with Bariatric Surgery, Weight / BMI, T2D and Diet (pFDR <0.05) was generated using Eulerr in R (version 6.1.1)[102].

DIABLO integration of omics datasets

To probe relationships between data sets we used Data Integration Analysis for Biomarker discovery using Latent cOmponents (DIABLO)[103], implemented through the mixOmics[104] package in R. DIABLO extends sparse generalized canonical correlation analysis (sGCCA)[105] to a classification framework. Resulting in a multi-omics integrative method that simultaneously identifies key variables correlated across different data types while discriminating between phenotypic groups.

Normalised datasets (gut microbiota species taxa, gut microbiota KEGG pathways (Levels 2-3) and quantified metabolites from urine, serum and faecal biofluids) were used with a full weighted design matrix where correlation was 0.1 between data matrices and 1 for the Y outcome, to result in a correlated and discriminant molecular signature[104]. To account for the repeated measures (pre/post procedure) experimental design, multilevel DIABLO (mDIABLO) models were constructed using within-subject variation matrices for each omics dataset[103, 106, 107]. Classification performance was assessed using the balanced error rate (BER) from cross-validation of samples, with BER = 0.5 x (false positive rate + false negative rate). BER scores range from 0-1, with a perfect classification model scoring 0, a random predictor 0.5 and a model with systematically incorrect predictions 1.

Gut Microbiota-Metabolome associations

Spearman's correlations between BMI, HbA1c and Microbiota-Metabolome datasets were generated in MATLAB. Corrected p values (pFDR) were used to select significant correlations. Significant

(pFDR<0.01) first order correlations to BMI and HbA1c and cross-correlations between these variables were displayed using Cytoscape 3.8.0.[108]. Correlations between gut microbiota, metabolite and dietary datasets were displayed using ComplexHeatmap in R[109]. Correlations with a pFDR value <0.05 are displayed, correlations with a pFDR <0.01 are highlighted. Hierarchical clustering of correlations was performed using Euclidean distances.

References

1. Qin J, *et al.* A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**, 55-60 (2012).

2. Ley RE, Turnbaugh PJ, Klein S, Gordon JI. Microbial ecology: human gut microbes associated with obesity. *Nature* **444**, 1022-1023 (2006).

3. Ridaura VK, *et al.* Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science* **341**, 1241214 (2013).

4. Vrieze A, *et al.* Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome. *Gastroenterology* **143**, 913-916 e917 (2012).

5. Heintz-Buschart A, Wilmes P. Human Gut Microbiome: Function Matters. *Trends Microbiol* **26**, 563-574 (2018).

6. Adams TD, *et al.* Health benefits of gastric bypass surgery after 6 years. *JAMA : the journal of the American Medical Association* **308**, 1122-1131 (2012).

7. Schauer PR, *et al.* Bariatric Surgery versus Intensive Medical Therapy for Diabetes - 5-Year Outcomes. *The New England journal of medicine* **376**, 641-651 (2017).

8. Cummings BP, *et al.* Vertical sleeve gastrectomy improves glucose and lipid metabolism and delays diabetes onset in UCD-T2DM rats. *Endocrinology* **153**, 3620-3632 (2012).

9. Liou AP, Paziuk M, Luevano JM, Jr., Machineni S, Turnbaugh PJ, Kaplan LM. Conserved shifts in the gut microbiota due to gastric bypass reduce host weight and adiposity. *Science translational medicine* **5**, 178ra141 (2013).

10. Liu R, *et al.* Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. *Nature medicine* **23**, 859-868 (2017).

11. Aron-Wisnewsky J, *et al.* Major microbiota dysbiosis in severe obesity: fate after bariatric surgery. *Gut* **68**, 70-82 (2019).

12. Turnbaugh PJ, *et al.* A core gut microbiome in obese and lean twins. *Nature* **457**, 480-484 (2009).

13. Chiuve SE, *et al.* Alternative dietary indices both strongly predict risk of chronic disease. *The Journal of nutrition* **142**, 1009-1018 (2012).

14. Perez-Pevida B, Escalada J, Miras AD, Fruhbeck G. Mechanisms Underlying Type 2 Diabetes Remission After Metabolic Surgery. *Frontiers in endocrinology* **10**, 641 (2019).

15. Stenberg E, Thorell A. Insulin resistance in bariatric surgery. *Current opinion in clinical nutrition and metabolic care* **23**, 255-261 (2020).

16. Perez-Arana GM, *et al.* Sleeve Gastrectomy and Roux-En-Y Gastric Bypass. Two Sculptors of the Pancreatic Islet. *J Clin Med* **10**, (2021).

17. Patti ME*, et al.* Heterogeneity of proliferative markers in pancreatic beta-cells of patients with severe hypoglycemia following Roux-en-Y gastric bypass. *Acta Diabetol* **54**, 737-747 (2017).

18. Miras AD, le Roux CW. Mechanisms underlying weight loss after bariatric surgery. *Nature reviews Gastroenterology & hepatology* **10**, 575-584 (2013).

19. Molinaro A, Wahlstrom A, Marschall HU. Role of Bile Acids in Metabolic Control. *Trends Endocrinol Metab* **29**, 31-41 (2018).

20. Penney NC, Kinross J, Newton RC, Purkayastha S. The role of bile acids in reducing the metabolic complications of obesity after bariatric surgery: a systematic review. *International journal of obesity* **39**, 1565-1574 (2015).

21. Larraufie P*, et al.* Important Role of the GLP-1 Axis for Glucose Homeostasis after Bariatric Surgery. *Cell Rep* **26**, 1399-1408 e1396 (2019).

22. Frezza EE, Chiriva-Internati M, Wachtel MS. Analysis of the results of sleeve gastrectomy for morbid obesity and the role of ghrelin. *Surgery today* **38**, 481-483 (2008).

23. Jin ZL, Liu W. Progress in treatment of type 2 diabetes by bariatric surgery. *World J Diabetes* **12**, 1187-1199 (2021).

24. Akalestou E, Miras AD, Rutter GA, le Roux CW. Mechanisms of weight loss after obesity surgery. *Endocrine reviews*, (2021).

25. Barberio MD*, et al.* Type 2 Diabetes Modifies Skeletal Muscle Gene Expression Response to Gastric Bypass Surgery. *Frontiers in endocrinology* **12**, 728593 (2021).

26. McTigue KM*, et al.* Comparing the 5-Year Diabetes Outcomes of Sleeve Gastrectomy and Gastric Bypass: The National Patient-Centered Clinical Research Network (PCORNet) Bariatric Study. *JAMA surgery* **155**, e200087 (2020).

27. Zhou W*, et al.* Longitudinal multi-omics of host-microbe dynamics in prediabetes. *Nature* **569**, 663-671 (2019).

28. Forslund K*, et al.* Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. *Nature* **528**, 262-266 (2015).

29. Wu H*, et al.* Metformin alters the gut microbiome of individuals with treatment-naive type 2 diabetes, contributing to the therapeutic effects of the drug. *Nature medicine* **23**, 850-858 (2017).

30. Wang TJ*, et al.* 2-Aminoadipic acid is a biomarker for diabetes risk. *The Journal of clinical investigation* **123**, 4309-4317 (2013).

31. Razquin C*, et al.* Lysine pathway metabolites and the risk of type 2 diabetes and cardiovascular disease in the PREDIMED study: results from two case-cohort studies. *Cardiovascular diabetology* **18**, 151 (2019).

32. Xu WY*, et al.* 2-Aminoadipic acid protects against obesity and diabetes. *J Endocrinol* **243**, 111-123 (2019).

33. Thorpe SR, Baynes JW. Role of the Maillard reaction in diabetes mellitus and diseases of aging. *Drugs Aging* **9**, 69-77 (1996).

34. Connick JH, Stone TW. The role of kynurenines in diabetes mellitus. *Med Hypotheses* **18**, 371-376 (1985).

35. Hirayama A*, et al.* Metabolic profiling reveals new serum biomarkers for differentiating diabetic nephropathy. *Analytical and bioanalytical chemistry* **404**, 3101-3109 (2012).

36. Liang X, Kaya A, Zhang Y, Le DT, Hua D, Gladyshev VN. Characterization of methionine oxidation and methionine sulfoxide reduction using methionine-rich cysteine-free proteins. *BMC Biochem* **13**, 21 (2012).

37. le Roux CW*, et al.* Gut hormone profiles following bariatric surgery favor an anorectic state, facilitate weight loss, and improve metabolic parameters. *Annals of surgery* **243**, 108-114 (2006).

38. Scholtz S*, et al.* Obese patients after gastric bypass surgery have lower brain-hedonic responses to food than after gastric banding. *Gut* **63**, 891-902 (2014).

39. Ahmed K, Penney N, Darzi A, Purkayastha S. Taste Changes after Bariatric Surgery: a Systematic Review. *Obesity surgery* **28**, 3321-3332 (2018).

40. Rennie KL, Coward A, Jebb SA. Estimating under-reporting of energy intake in dietary surveys using an individualised method. *The British journal of nutrition* **97**, 1169-1176 (2007).

41. Donaldson GP, Lee SM, Mazmanian SK. Gut biogeography of the bacterial microbiota. *Nature reviews Microbiology* **14**, 20-32 (2016).

42. Zoetendal EG*, et al.* The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *ISME J* **6**, 1415-1426 (2012).

43. Suzuki TA, Nachman MW. Spatial Heterogeneity of Gut Microbial Composition along the Gastrointestinal Tract in Natural Populations of House Mice. *PloS one* **11**, e0163720 (2016).

44. Miquel S*, et al.* Ecology and metabolism of the beneficial intestinal commensal bacterium Faecalibacterium prausnitzii. *Gut microbes* **5**, 146-151 (2014).

45. Ley RE, Backhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. Obesity alters gut microbial ecology. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 11070-11075 (2005).

46. Duncan SH*, et al.* Human colonic microbiota associated with diet, obesity and weight loss. *International journal of obesity* **32**, 1720-1724 (2008).

47. Schwiertz A*, et al.* Microbiota and SCFA in lean and overweight healthy subjects. *Obesity* **18**, 190-195 (2010).

48. Sze MA, Schloss PD. Looking for a Signal in the Noise: Revisiting Obesity and the Microbiome. *mBio* **7**, (2016).

49. Elliott P*, et al.* Urinary metabolic signatures of human adiposity. *Science translational medicine* **7**, 285ra262 (2015).

50. Li JV*, et al.* Metabolic surgery profoundly influences gut microbial-host metabolic cross-talk. *Gut* **60**, 1214-1223 (2011).

51. Pallister T*, et al.* Hippurate as a metabolomic marker of gut microbiome diversity: Modulation by diet and relationship to metabolic syndrome. *Sci Rep* **7**, 13670 (2017).

52. Tai ES*, et al.* Insulin resistance is associated with a metabolic profile of altered protein metabolism in Chinese and Asian-Indian men. *Diabetologia* **53**, 757-767 (2010).

53. Wang TJ*, et al.* Metabolite profiles and the risk of developing diabetes. *Nature medicine* **17**, 448-453 (2011).

54. Soeters PB, Fischer JE. Insulin, glucagon, aminoacid imbalance, and hepatic encephalopathy. *Lancet* **2**, 880-882 (1976).

55. Libert DM, Nowacki AS, Natowicz MR. Metabolomic analysis of obesity, metabolic syndrome, and type 2 diabetes: amino acid and acylcarnitine levels change along a spectrum of metabolic wellness. *PeerJ* **6**, e5410 (2018).

56. White PJ, Newgard CB. Branched-chain amino acids in disease. *Science* **363**, 582-583 (2019).

57. Tremblay F*, et al.* Identification of IRS-1 Ser-1101 as a target of S6K1 in nutrient- and obesity-induced insulin resistance. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 14056-14061 (2007).

58. Mardinoglu A*, et al.* The gut microbiota modulates host amino acid and glutathione metabolism in mice. *Molecular systems biology* **11**, 834 (2015).

59. Pedersen HK*, et al.* Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature* **535**, 376-381 (2016).

60. Dai ZL, Wu G, Zhu WY. Amino acid metabolism in intestinal bacteria: links between gut ecology and host health. *Front Biosci (Landmark Ed)* **16**, 1768-1786 (2011).

61. Hoyles L*, et al.* Metabolic retroconversion of trimethylamine N-oxide and the gut microbiota. *Microbiome* **6**, 73 (2018).

62. Koeth RA*, et al.* Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nature medicine* **19**, 576-585 (2013).

63. Mueller DM*, et al.* Plasma levels of trimethylamine-N-oxide are confounded by impaired kidney function and poor metabolic control. *Atherosclerosis* **243**, 638-644 (2015).

64. Barton W*, et al.* The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level. *Gut* **67**, 625-633 (2018).

65. Holmes E*, et al.* Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature* **453**, 396-400 (2008).

66. Dumas ME*, et al.* Microbial-Host Co-metabolites Are Prodromal Markers Predicting Phenotypic Heterogeneity in Behavior, Obesity, and Impaired Glucose Tolerance. *Cell Rep* **20**, 136-148 (2017).

67. Hardison WG, Grundy SM. Effect of bile acid conjugation pattern on bile acid metabolism in normal humans. *Gastroenterology* **84**, 617-620 (1983).

68. Swann JR*, et al.* Systemic gut microbial modulation of bile acid metabolism in host tissue compartments. *Proceedings of the National Academy of Sciences of the United States of America* **108 Suppl 1**, 4523-4530 (2011).

69. Zheng X*, et al.* Hyocholic acid species as novel biomarkers for metabolic disorders. *Nature communications* **12**, 1487 (2021).

70. Zheng X*, et al.* Hyocholic acid species improve glucose homeostasis through a distinct TGR5 and FXR signaling mechanism. *Cell metabolism* **33**, 791-803 e797 (2021).

71. Perry RJ*, et al.* Acetate mediates a microbiome-brain-beta-cell axis to promote metabolic syndrome. *Nature* **534**, 213-217 (2016).

72. Lin HV*, et al.* Butyrate and propionate protect against diet-induced obesity and regulate gut hormones via free fatty acid receptor 3-independent mechanisms. *PloS one* **7**, e35240 (2012).

73. Windey K, De Preter V, Verbeke K. Relevance of protein fermentation to gut health. *Mol Nutr Food Res* **56**, 184-196 (2012).

74. Hughes R, Magee EA, Bingham S. Protein degradation in the large intestine: relevance to colorectal cancer. *Curr Issues Intest Microbiol* **1**, 51-58 (2000).

75. Li JV*, et al.* Experimental bariatric surgery in rats generates a cytotoxic chemical environment in the gut contents. *Frontiers in microbiology* **2**, 183 (2011).

76. Sun CY, Hsu HH, Wu MS. p-Cresol sulfate and indoxyl sulfate induce similar cellular inflammatory gene expressions in cultured proximal renal tubular cells. *Nephrol Dial Transplant* **28**, 70-78 (2013).

77. Gratton J*, et al.* Optimized Sample Handling Strategy for Metabolic Profiling of Human Feces. *Analytical chemistry* **88**, 4661-4668 (2016).

78. Dona AC*, et al.* Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping. *Analytical chemistry* **86**, 9887-9894 (2014).

79. Veselkov KA*, et al.* Optimized preprocessing of ultra-performance liquid chromatography/mass spectrometry urinary metabolic profiles for improved information recovery. *Analytical chemistry* **83**, 5864-5872 (2011).

80. Sarafian MH*, et al.* Bile acid profiling and quantification in biofluids using ultra-performance liquid chromatography tandem mass spectrometry. *Analytical chemistry* **87**, 9662-9670 (2015).

81. Moreau NM*, et al.* Simultaneous measurement of plasma concentrations and 13C-enrichment of short-chain fatty acids, lactic acid and ketone bodies by gas chromatography coupled to mass spectrometry. *Journal of chromatography B, Analytical technologies in the biomedical and life sciences* **784**, 395-403 (2003).

82. Siskos AP*, et al.* Interlaboratory Reproducibility of a Targeted Metabolomics Platform for Analysis of Human Serum and Plasma. *Analytical chemistry* **89**, 656-665 (2017).

83. Xu H*, et al.* FastUniq: a fast de novo duplicates removal tool for paired short reads. *PloS one* **7**, e52249 (2012).

84. Hoyles L*, et al.* Molecular phenomics and metagenomics of hepatic steatosis in non-diabetic obese women. *Nature medicine* **24**, 1070-1080 (2018).

85. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*, (2013).

86. Truong DT*, et al.* MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat Methods* **12**, 902-903 (2015).

87. Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C. Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat Methods* **9**, 811-814 (2012).

88. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. *Genome Res* **27**, 824-834 (2017).

89. Zhu W, Lomsadze A, Borodovsky M. Ab initio gene identification in metagenomic sequences. *Nucleic Acids Res* **38**, e132 (2010).

90. Besemer J, Borodovsky M. Heuristic approach to deriving models for gene finding. *Nucleic Acids Res* **27**, 3911-3920 (1999).

91. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460-2461 (2010).

92. Le Chatelier E*, et al.* Richness of human gut microbiome correlates with metabolic markers. *Nature* **500**, 541-546 (2013).

93. Huerta-Cepas J*, et al.* Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Mol Biol Evol* **34**, 2115-2122 (2017).

94. Qin N*, et al.* Alterations of the human gut microbiome in liver cirrhosis. *Nature* **513**, 59-64 (2014).

95. Warrack BM*, et al.* Normalization strategies for metabonomic analysis of urine samples. *Journal of chromatography B, Analytical technologies in the biomedical and life sciences* **877**, 547-552 (2009).

96. Asnicar F, Weingart G, Tickle TL, Huttenhower C, Segata N. Compact graphical representation of phylogenetic data and metadata with GraPhlAn. *PeerJ* **3**, e1029 (2015).

97. Posma JM*, et al.* Optimized Phenotypic Biomarker Discovery and Confounder Elimination via Covariate-Adjusted Projection to Latent Structures from Metabolic Spectroscopy Data. *Journal of proteome research* **17**, 1586-1595 (2018).

98. Garcia-Perez I*, et al.* Objective assessment of dietary patterns by use of metabolic phenotyping: a randomised, controlled, crossover trial. *The lancet Diabetes & endocrinology* **5**, 184-195 (2017).

99. Oksanen J*, et al.* vegan: Community Ecology Package. *R package* **version 2.5-6**, (2019).

100. Posma JM*, et al.* Subset optimization by reference matching (STORM): an optimized statistical approach for recovery of metabolic biomarker structural information from 1H NMR spectra of biofluids. *Analytical chemistry* **84**, 10694-10701 (2012).

101. Cloarec O*, et al.* Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H NMR data sets. *Analytical chemistry* **77**, 1282-1289 (2005).

102. Larsson J. eulerr: Area-Proportional Euler and Venn Diagrams with Ellipses. R package version 6.1.1. (2021).

103. Singh A*, et al.* DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics* **35**, 3055-3062 (2019).

104. Rohart F, Gautier B, Singh A, Le Cao KA. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol* **13**, e1005752 (2017).

105. Tenenhaus A, Philippe C, Guillemot V, Le Cao KA, Grill J, Frouin V. Variable selection for generalized canonical correlation analysis. *Biostatistics* **15**, 569-583 (2014).

106. Liquet B, Le Cao KA, Hocini H, Thiebaut R. A novel approach for biomarker selection and the integration of repeated measures experiments from two assays. *BMC bioinformatics* **13**, 325 (2012).

107. Westerhuis JA, van Velzen EJ, Hoefsloot HC, Smilde AK. Multivariate paired data analysis: multilevel PLSDA versus OPLSDA. *Metabolomics* **6**, 119-128 (2010).

108. Shannon P*, et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504 (2003).

109. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847-2849 (2016).

Acknowledgments:

This research was funded by the Diabetes Research and Wellness Foundation through the Sutherland-Earl Clinical Research Fellowship 2015 awarded to NP.

The research was supported by the National Institute for Health Research (NIHR) Imperial Biomedical Research Centre (BRC) and Clinical Research Network (CRN). The views expressed are those of the authors and not necessarily those of the NHS, NIHR or UK Department of Health and Social Care.

We are grateful to the Imperial Weight Centre for facilitating this research.

This work used the computing resources of the UK MEDical BIOinformatics partnership – aggregation, integration, visualisation and analysis of large, complex data (UK Med-Bio) – which was supported by the Medical Research Council (grant number MR/L01632X/1).

IGP is supported by a NIHR Career Development Research Fellowship (NIHR-CDF-2017-10-032). JMP is supported by a Rutherford Fund Fellowship at Health Data Research (HDR) UK (MR/S004033/1). GF is supported by an NIHR Senior Investigator award. LH was in receipt of an MRC Intermediate Research Fellowship in Data Science (MR/L01632X/1, UK MED-BIO). EH is supported by the Department of Jobs, Tourism, Science and Innovation, Government of Western Australian through the Premier's Science Fellowship Program and by the Australian Research Council Laureate Fellowship Scheme.

Author contributions:

Conceptualisation: NP, EH, JM, AD. Funding acquisition: NP, SP. Data curation: NP, DY, BG. Microbiome Investigation: NP, LH, JM. Metabolite Investigation: NP, IGP, EH. Microbiome Analysis: LH, NP. Metabolite Analysis: NP, IGP, EH. Dietary Analysis: AK, NP, GF. Software: JMP, GF. Supervision: EH, SP, JM, GF, AD. Resources: EH, SP, JM, GF, AD. Writing – Original Draft: NP. Writing – Review & Editing: All authors.

Declaration of interests:

The authors declare no competing interests.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementalInformationFINAL.pdf